Bugni, Federico A.; Li, Jia; Li, Qiyuan

Article

Permutation-based tests for discontinuities in event studies

Quantitative Economics

Provided in Cooperation with:
The Econometric Society

Suggested Citation: Bugni, Federico A.; Li, Jia; Li, Qiyuan (2023) : Permutation-based tests for discontinuities in event studies, Quantitative Economics, ISSN 1759-7331, The Econometric Society, New Haven, CT, Vol. 14, Iss. 1, pp. 37-70, https://doi.org/10.3982/QE1775

Quantitative Economics 14 (2023), 37–70 1759-7331/20230037

Permutation-based tests for discontinuities in event studies

Federico A. Bugni
Department of Economics, Northwestern University

Jia Li
School of Economics, Singapore Management University

Qiyuan Li
School of Economics, Singapore Management University

We propose using a permutation test to detect discontinuities in an underlying economic model at a known cutoff point. Relative to the existing literature, we show that this test is well suited for event studies based on time-series data. The test statistic measures the distance between the empirical distribution functions of observed data in two local subsamples on the two sides of the cutoff. Critical values are computed via a standard permutation algorithm. Under a high-level condition that the observed data can be coupled by a collection of conditionally independent variables, we establish the asymptotic validity of the permutation test, allowing the sizes of the local subsamples to be either be fixed or grow to infinity. In the latter case, we also establish that the permutation test is consistent. We demonstrate that our high-level condition can be verified in a broad range of problems in the infill asymptotic time-series setting, which justifies using the permutation test to detect jumps in economic variables such as volatility, trading activity, and liquidity. These potential applications are illustrated in an empirical case study for selected FOMC announcements during the ongoing COVID-19 pandemic.

Keywords. Event study, infill asymptotics, jump, permutation tests, randomization tests, semimartingale.

JEL classification. C12, C14, C22, C32.

1. Introduction

Many econometric problems can be expressed in terms of the continuity or the discontinuity of certain component in the underlying economic model. In an influential paper, Chow (1960) tested the temporal stability in the demand for automobiles, and subsequently stimulated a large literature on structural breaks in time-series analysis;

Federico A. Bugni: federico.bugni@northwestern.edu
Jia Li: jiali@smu.edu.sg
Qiyuan Li: qyli.2019@phdecons.smu.edu.sg
We thank the two anonymous referees for comments and suggestions that have greatly improved the manuscript. A replication file is posted (Bugni, Li, and Li (2023)).

 https://doi.org/10.3982/QE1775

see, for example, Andrews (1993), Stock (1994), Bai and Perron (1998), and many references therein. In microeconometrics, the regression discontinuity design (RDD) has been extensively used for causal inference. This literature identifies and estimates an average treatment effect by evaluating discontinuities of conditional expectation functions of outcome and treatment variables at a cutoff point of the running variable; see Imbens and Lemieux (2008) and Lee and Lemieux (2010) for comprehensive reviews.[1] Meanwhile, a more recent high-frequency financial econometrics literature has been devoted to studying discontinuities, or jumps, in various financial time series (e.g., price, volatility, trading activity, etc.). The high-frequency jump literature is pioneered by Barndorff-Nielsen and Shephard (2006), who propose the first nonparametric test for asset price jumps using high-frequency data in an infill asymptotic setting. More recently, Bollerslev, Li, and Xue (2018) study the jumps of volatility and trading intensity in high-frequency jump regressions (Li, Todorov, and Tauchen (2017)) that closely resemble the classical RDD.

Although these strands of literature involve apparently different terminology and technical tools, they share a common theme: The econometric goal is to learn about differences in the data generating processes between two subsamples separated by the cutoff. Imbens and Kalyanaraman (2011) emphasize that these subsamples should be "local" to the cutoff point, which is quite natural given the nonparametric nature of discontinuity inference (Hahn, Todd, and Van der Klaauw (2001)). The issue under study is thus a local version of the classical two-sample problem. Correspondingly, the related inference is often carried out using nonparametric two-sample t-tests, which are based on kernel regressions in the RDD (Hahn, Todd, and Van der Klaauw (2001), Imbens and Kalyanaraman (2011), Calonico, Cattaneo, and Titiunik (2014)), or in the same spirit, spot high-frequency estimators (Foster and Nelson (1996), Comte and Renault (1998), Jacod and Protter (2012), Li, Todorov, and Tauchen (2017), Bollerslev, Li, and Xue (2018)) in the infill time-series setting.

In an ideal scenario in which the subsamples separated by the cutoff are i.i.d. and independent of each other, the permutation test is an excellent tool to detect differences in their distributions. In particular, standard results for randomization inference (Lehmann and Romano (2005, Chapter 15.2)) indicate that a permutation test implemented with any arbitrary test statistic is finite-sample valid under these conditions. The recent literature has investigated the properties of permutation tests to detect differences between two samples under less ideal conditions. One example is Canay and Kamat (2017), who consider an RDD and show that permutation-based inference is asymptotically valid to detect discontinuities in the distribution of the baseline covariates at the cutoff. These authors implement their test with a finite number of observations that are located closest to the cutoff, effectively forcing them to concentrate on a small neighborhood of the cutoff as the sample size grows. In the same spirit, Cattaneo, Titiunik, and Vazquez-Bare (2017) propose using permutation-based inference to detect discontinuities at the cutoff under the "local randomization framework" introduced in

[1]Coincidentally, the RDD was first proposed by Thistlethwaite and Campbell (1960) around the same time as the Chow test.

Cattaneo, Frandsen, and Titiunik (2015). Outside of the RDD literature, Chung and Romano (2013) and DiCiccio and Romano (2017) investigate the asymptotic properties of permutation-based inference to test for differences in specific distributional features of two samples, such as the mean or the correlation coefficient. It is important to note that all of the references mentioned in this paragraph presume cross-sectional data.

In the context of time-series applications, there is an active literature on change-point tests implemented via permutations. This approach was first suggested by Antoch and Hušková (2001) and later pursued by other authors; see Hušková (2004) or Horváth and Rice (2014) for surveys of this literature. While most of this literature imposes independent errors, some allow for limited forms of weak dependence (Kirch and Steinebach (2006), Kirch (2007), and Jentsch and Pauly (2015)). In contrast, our econometric setting accommodates essentially unrestricted persistence and nonstationarity in the underlying state processes (e.g., volatility), which better suits our interest on their dynamics over short time windows around economic news events. In the context of machine-learning methods, Chernozhukov, Wuthrich, and Zhu (2018) propose using permutations to implement conformal inference that allows for time-series data.

Set against this background, our main goal in this paper is to establish a general theory for permutation-based discontinuity tests, with a special emphasis on event studies based on time-series data. To capture the "local" nature of this problem, we adopt an infill asymptotic framework, under which the inference concentrates on observations "close" to the event time. Specifically, we consider the Cramér–von Mises test statistic formed as the squared L_2 distance between the empirical cumulative distribution functions for the two local subsamples near the cutoff, and compute the critical value via a standard permutation algorithm. As explained earlier, if the data were i.i.d., the behavior of this permutation test would follow directly from standard results for randomization inference. This "off-the-shelf" theory, however, is not applicable here because time-series data observed in a short event window can be serially highly dependent.

The main theoretical contribution of the present paper is to establish the asymptotic validity of the permutation test in this nonstandard setting. The theory has two components. The first is a new general result for permutation test. Specifically, we link the (feasible) permutation test formed using the original data with an infeasible test constructed in a "coupling" problem that involves conditionally i.i.d. coupling variables. Since the latter resembles the classical two-sample problem, the infeasible test controls size exactly under the coupling null hypothesis (i.e., coupling variables in the two subsamples are homogeneous), and is consistent under the complementary alternative hypothesis. Under a proper notion of coupling, which is customized for the permutation test, we show that the feasible test inherits the same asymptotic rejection properties from the infeasible one. Since this result is of independent theoretical interest that is well beyond our subsequent analysis in the infill time-series setting, we frame the theory under general high-level conditions so as to facilitate other types of applications.

The second component of our analysis pertains to specializing the general result to the infill time-series setting designed for event-study applications. The event-study framework is particularly relevant for studying macroeconomic and financial shocks, including monetary shocks triggered by FOMC announcements (Cochrane and Piazzesi

(2002), Nakamura and Steinsson (2018a)), or "natural disasters" such as the ongoing COVID-19 pandemic. Following Li and Xiu (2016) and Bollerslev, Li, and Xue (2018), we model observed data using a general state-space framework, in which the observations are discretely sampled from a latent state process "contaminated" by random disturbances. This model has been used to model variables such as asset returns, trading volume, duration, and bid-ask spread, and readily accommodates both continuously and discretely valued variables. Under this state-space model, the temporal discontinuity in the data's distribution is mainly driven by the jump of the latent state process (e.g., asset volatility, trading intensity, and propensity of informed trading), which can be detected by the permutation test. Under easy-to-verify primitive conditions, we construct coupling variables and apply the aforementioned general theory to establish the permutation test's asymptotic validity.

We recognize two advantages of the proposed permutation test in comparison with the standard approach based on the nonparametric "spot" estimation of the underlying state process. First, the permutation test attains asymptotic size control even if the number of observations in each subsample is *fixed*.[2] This remarkable property is reminiscent of the finite-sample exactness of the permutation test in the classical two-sample problem for i.i.d. data. In contrast, the nonparametric estimation approach works in a fundamentally different way, as it relies on the asymptotic (mixed) normality of the estimator, which in turn requires the sizes of the local subsamples to grow to infinity. In empirical applications, however, it is often desirable to use a short time window, either to reduce the effect of confounding factors in the background, or simply because of the lack of observations soon after the occurrence of the economic event (say, in a real-time research situation). Not surprisingly, the conventional inference based on asymptotic Gaussianity often results in large size distortions in this "small sample" scenario, as we demonstrate concretely in a realistically calibrated Monte Carlo experiment (see Section 3). Meanwhile, the permutation test exhibits much more robust size control in finite samples.

The second advantage of the permutation test is its versatility: The same test can be applied in many different empirical contexts without any modification. On the other hand, the nonparametric estimation approach often relies on specific features of the problem, and needs to be designed on a case-by-case basis. Therefore, the proposed permutation test may be particularly attractive in new empirical environments for which tests based on the conventional approach are not yet developed or not yet well understood. In Section 2.2, we illustrate this point more concretely in the context of testing for volatility jumps. In that case, the standard approach relies crucially on the assumption that the price shocks are Brownian in its design of the spot volatility estimator and the associated t-statistic, and it cannot be adapted easily to accommodate a more general setting with Lévy-driven shocks.[3] The permutation test, on the other hand, is valid even in the latter more general setting.

[2]For a similar type of results in the context of RDD; see Cattaneo, Frandsen, and Titiunik (2015), Cattaneo, Titiunik, and Vazquez-Bare (2017), Canay and Kamat (2017), and Bugni and Canay (2021).

[3]As explained by Barndorff-Nielsen and Shephard (2001), these more general processes offer the possibility of capturing important deviations from Brownian shocks and for flexible modeling of dependence structures. However, to the best of our knowledge, the estimation and inference of the spot volatility (i.e., the

That being said, we stress that the proposed permutation test is a complement, rather than substitute, for the conventional nonparametric estimation method, because it has two limitations. One is that the permutation test focuses exclusively on hypothesis testing, without producing a point estimate for the jump of the state process (e.g., volatility) of interest, whereas the estimate is a byproduct of the conventional approach. In addition, the proposed permutation test is purely nonparametric and it does not exploit any parametric structure that one may be willing to impose. It is therefore conceivable that in certain semiparametric settings, more efficient tests may be designed to exploit a priori model restrictions. Put differently, the aforementioned versatility of the permutation test may come with an efficiency cost. A better understanding about the robustness-efficiency tradeoff might be an interesting topic for future research.

In an empirical illustration, we apply the permutation test to a recent sample of high-frequency intraday returns of the SPY ETF for the S&P 500 index. Specifically, we focus on four important FOMC announcements during the ongoing COVID-19 pandemic, and test whether each announcement induces discontinuities in volatility, trading activity, and two measures of market illiquidity. We document robust empirical evidence for discontinuities in volatility and trading activity. We also find evidence for announcement-induced discontinuity in transaction cost (measured by bid-ask spread), but not in market impact (gauged by Amihud's measure). This application highlights one of the main advantages of the proposed test; namely, it is applicable for a broad variety of high-frequency observations modeled in distinct ways, which is unlike, for example, the conventional t-test designed specifically for testing volatility jumps in the Brownian setting.

The rest of the paper is organized as follows. We present the asymptotic theory for the permutation test in Section 2. Section 3 reports the test's finite-sample performance in Monte Carlo experiments, and Section 4 presents the empirical illustration. Section 5 concludes. Appendix A contains all proofs and Appendix B discusses an extension of our methods.

Notation We use $\|x\|$ to denote the Euclidean norm of a vector x. For any real number a, we use $\lceil a \rceil$ to denote the smallest integer that is larger than a. For any constant $p \geq 1$, $\|\cdot\|_p$ denotes the L_p norm for random variables. For two real sequences a_n and b_n, we write $a_n \asymp b_n$ if $a_n/C \leq b_n \leq C a_n$ for some finite constant $C \geq 1$.

2. Theory

2.1 *A general result for the asymptotic validity of permutation tests*

We first prove a new result that is broadly useful for establishing the asymptotic validity of permutation tests. Because of its independent theoretical interest, we develop the theory under high-level conditions. In Section 2.2, below, we shall specialize this general result in event-study applications under a more specific infill time-series setting, for which the existing theory on permutation tests is not applicable.

scaling process) in the non-Brownian case remains to be an open question in the literature. There is some limited work on the inference of integrated volatility functionals for the non-Brownian case (see Todorov and Tauchen (2012)), which demonstrates various distinct complications in the non-Brownian setting.

Consider an array $(Y_{n,i})_{i\in\mathcal{I}_n}$ of $\mathbb{R}$-valued observed variables defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, which may be either "raw" data or preliminary estimators. Our econometric goal is to decide whether two subsamples $(Y_{n,i})_{i\in\mathcal{I}_{1,n}}$ and $(Y_{n,i})_{i\in\mathcal{I}_{2,n}}$ have "significantly" different distributions, where $(\mathcal{I}_{1,n}, \mathcal{I}_{2,n})$ is a partition of $\mathcal{I}_n \subseteq \mathbb{Z}$. For ease of exposition, we assume that $\mathcal{I}_{1,n}$ and $\mathcal{I}_{2,n}$ contain the same number of observations, denoted by k_n.[4] We stress from the outset that k_n may either be fixed or grow to infinity in the subsequent analysis. As such, our analysis speaks to not only the classical finite-sample analysis of permutation tests, but also the large sample analysis routinely used in econometrics.

To implement the test, we first estimate the empirical cumulative distribution functions (CDF) for the two subsamples using

$$\widehat{F}_{j,n}(x) \equiv \frac{1}{k_n} \sum_{i\in\mathcal{I}_{j,n}} 1\{Y_{n,i} \le x\}, \quad j \in \{1, 2\}.$$

We then measure their difference via the Cramér–von Mises statistic, given by

$$\widehat{T}_n \equiv \frac{1}{2k_n} \sum_{i\in\mathcal{I}_n} \big(\widehat{F}_{1,n}(Y_{n,i}) - \widehat{F}_{2,n}(Y_{n,i})\big)^2. \tag{1}$$

For a significance level $\alpha \in (0, 1)$, we compute the critical value via a standard permutation algorithm as in Lehmann and Romano (2005, p. 633), which we specify in Algorithm 1 below. We use π to denote a permutation of the elements of $\mathcal{I}_n$, that is, a bijective mapping from $\mathcal{I}_n$ to itself. Let G_n denote the collection of all possible permutations of $\mathcal{I}_n$, with M_n being its cardinality.

Algorithm 1. Step 1. For each permutation $\pi \in G_n$, compute the permuted test statistic $\widehat{T}_n(\pi)$ as $\widehat{T}_n$, but with $(Y_{n,i})_{i\in\mathcal{I}_n}$ replaced by $(Y_{n,\pi(i)})_{i\in\mathcal{I}_n}$.

Step 2. Order $\{\widehat{T}_n(\pi) : \pi \in G_n\}$ as $\widehat{T}_n^{(1)} \le \cdots \le \widehat{T}_n^{(M_n)}$. Set $\widehat{T}_n^* = \widehat{T}_n^{(k)}$ for $k = \lceil M_n(1-\alpha)\rceil$.

Step 3. If $\widehat{T}_n > \widehat{T}_n^*$, reject the null hypothesis. If $\widehat{T}_n < \widehat{T}_n^*$, do not reject the null hypothesis. If $\widehat{T}_n = \widehat{T}_n^*$, reject the null hypothesis with probability $\hat{p}_n \equiv (M_n\alpha - \widehat{M}_n^+)/\widehat{M}_n^0$, where $\widehat{M}_n^+$ and $\widehat{M}_n^0$ are the cardinalities of $\{j : \widehat{T}_n^{(j)} > \widehat{T}_n^*\}$ and $\{j : \widehat{T}_n^{(j)} = \widehat{T}_n^*\}$, respectively. The resulting test then rejects according to the critical function $\hat{\phi}_n \equiv 1\{\widehat{T}_n > \widehat{T}_n^*\} + \hat{p}_n 1\{\widehat{T}_n = \widehat{T}_n^*\}$.

Remark 2.1. The test specified in Algorithm 1 is a randomized test and has a random outcome when $\widehat{T}_n = \widehat{T}_n^*$ (i.e., it rejects the null hypothesis with probability $\hat{p}_n$ independent of the data); see Section 3.1 in Lehmann and Romano (2005) for details about randomized tests. One can construct a nonrandomized (and more conservative) version by replacing $\hat{p}_n$ with zero. Also, in practice, M_n may be too large to consider G_n in its entirety. In such cases, we could replace G_n with a random subset of it, denoted by $\widehat{G}_n$, and composed of the identity permutation and an i.i.d. sample of permutations in G_n. All of the formal results in this paper apply if we use $\widehat{G}_n$ instead of G_n in Algorithm 1.

[4] All of our results can be easily extended to the case when $\mathcal{I}_{1,n}$ and $\mathcal{I}_{2,n}$ have different sizes, but with the same order of magnitude.

REMARK 2.2. In this paper, we use the Cramér–von Mises statistic in (1) for concreteness and simplicity of exposition. However, Appendix B of this paper shows that our main results extend well beyond this particular statistic. In particular, the asymptotic validity of the permutation test extends to any other rank statistic, that is, a statistic that only depends on the rank of the observations. Furthermore, our consistency result applies to any other rank statistic under mild regularity conditions. For example, both of these results hold if we replace (1) with the Kolmogorov–Smirnov statistic, given by

$$\widehat{T}_n \equiv \max_{i\in\mathcal{I}_n} \big|\widehat{F}_{1,n}(Y_{n,i}) - \widehat{F}_{2,n}(Y_{n,i})\big|.$$

For details, see Appendix B.

If the data $(Y_{n,i})_{i\in\mathcal{I}_n}$ are i.i.d., then the null hypothesis of the classical two-sample problem holds, and Lehmann and Romano (2005, Theorem 15.2.1) implies that the aforementioned permutation test has *exact* size control in finite samples. This is a remarkable property of the permutation test, as it holds without requiring any specific distributional assumptions on the data. In contrast to the classical two-sample problem, however, we shall not assume that the data are independent, or even "weakly" dependent (e.g., mixing). As mentioned in the Introduction, the main goal of this paper is to study the permutation test for time-series data observed within a short event window (say, a few days or hours), which can be serially highly dependent in practice. Our key theoretical insight is that the permutation test is still asymptotically valid if the data $(Y_{n,i})_{i\in\mathcal{I}_n}$ can be approximated, or "coupled," by another collection of variables that are conditionally independent, as formalized by the following assumption.

ASSUMPTION 2.1. *There exists a collection of variables* $(U_{n,i})_{i\in\mathcal{I}_n}$ *such that the following conditions hold for a sequence* $(\mathcal{G}_n)_{n\geq 1}$ *of* σ*-fields*: (*i*) *for each* $n \geq 1$, *the variables* $(U_{n,i})_{i\in\mathcal{I}_n}$ *are* $\mathcal{G}_n$*-conditionally independent, and* $U_{n,i}$ *has the same* $\mathcal{G}_n$*-conditional distribution as* $U_{n,j}$ *if* i, j *belong to the same subsample* (*i.e.*, $\mathcal{I}_{1,n}$ *or* $\mathcal{I}_{2,n}$); (*ii*) *for any real sequence* $\eta_n = o(1)$, *we have* $\sup_{x\in\mathbb{R}} \mathbb{P}(|U_{n,i} - x| \leq \eta_n|\mathcal{G}_n) = O_p(\eta_n)$; (*iii*) $\max_{i\in\mathcal{I}_n} |\widetilde{Y}_{n,i} - U_{n,i}| = o_p(k_n^{-2})$, *where* $(\widetilde{Y}_{n,i})_{i\in\mathcal{I}_n}$ *is an identical copy of* $(Y_{n,i})_{i\in\mathcal{I}_n}$ *in* $\mathcal{G}_n$*-conditional distribution*.

Assumption 2.1 lays out the high-level structure for bridging our analysis with the classical theory on permutation tests, which we carry out in Theorem 2.1 below. Condition (i) sets up the "coupling" problem, which corresponds to a conditional version of the classical two-sample problem, treating the $(U_{n,i})_{i\in\mathcal{I}_{1,n}}$ and $(U_{n,i})_{i\in\mathcal{I}_{2,n}}$ variables as "data." In part (a) of Theorem 2.1, we consider the situation in which both subsamples have the same conditional distribution. In this case, our coupling variables $(U_{n,i})_{i\in\mathcal{I}_n}$ give rise to an infeasible permutation test that can be analyzed as a classical two-sample problem. In particular, this infeasible permutation test attains the exact finite-sample size under our conditions.

This infeasible test, however, only plays an auxiliary role in our analysis, because our interest is on the feasible test $\hat{\phi}_n$ formed using the original $(Y_{n,i})_{i\in\mathcal{I}_n}$ data. Therefore,

a key component of our theoretical argument in Theorem 2.1 is to show that the feasible test for the original data inherits asymptotically the same rejection properties from the infeasible test. Conditions (ii) and (iii) in Assumption 2.1 are introduced for this purpose. Specifically, condition (ii) requires the variable $U_{n,i}$ to be nondegenerate, in the sense that its conditional probability mass within any small $[x-\eta, x+\eta]$ interval is of order $O(\eta)$ in probability.[5] Condition (iii) specifies the requisite approximation accuracy of the coupling variables. We note that this condition is easier to hold when k_n is smaller, because the joint coupling requirement would involve a smaller number of variables and the $o_p(k_n^{-2})$ error bound is easier to attain. This condition can be verified under more primitive conditions pertaining to the smoothness of underlying processes and an upper bound on the growth rate of k_n, as detailed in Section 2.2.[6]

THEOREM 2.1. *Under Assumption 2.1, the following statements hold for the permutation test $\hat{\phi}_n$ described in Algorithm 1:*

(a) If the variables $(U_{n,i})_{i\in\mathcal{I}_n}$ have the same $\mathcal{G}_n$-conditional distribution, we have $\mathbb{E}[\hat{\phi}_n] \to \alpha$.

(b) Let $Q_{j,n}(\cdot)$ denote the $\mathcal{G}_n$-conditional distribution function of $U_{n,i}$ for $i \in \mathcal{I}_{j,n}$ and $j \in \{1,2\}$, and $\overline{Q}_n = (Q_{1,n}+Q_{2,n})/2$. If $k_n \to \infty$ and $\mathbb{P}(\int (Q_{1,n}(x)-Q_{2,n}(x))^2 d\overline{Q}_n(x) > \delta_n) \to 1$ for any real sequence $\delta_n = o(1)$, we have $\mathbb{E}[\hat{\phi}_n] \to 1$.

Theorem 2.1 characterizes the asymptotic rejection probabilities of the feasible test $\hat{\phi}_n$ under the null and alternative hypotheses of the two-sample problem for the coupling variables. Part (a) pertains to the situation in which the two subsamples of coupling variables, $(U_{n,i})_{i\in\mathcal{I}_{1,n}}$ and $(U_{n,i})_{i\in\mathcal{I}_{2n}}$, have the same conditional distribution, which corresponds to the null hypothesis. In this case, the theorem shows that the asymptotic rejection probability of the feasible test is equal to the nominal level α. It is relevant to note that this result holds whether k_n is fixed or divergent. This property is clearly reminiscent of the permutation test's finite-sample exactness in the classical setting.

Part (b) of Theorem 2.1 concerns the power of the feasible test $\hat{\phi}_n$. It shows that the feasible test rejects with probability approaching one when the conditional distributions of the two coupling subsamples, $Q_{1,n}$ and $Q_{2,n}$, are different, in the sense that their "distance" measured by $\int (Q_{1,n}(x)-Q_{2,n}(x))^2 d\overline{Q}_n(x)$ is asymptotically nondegenerate, where the mixture distribution $\overline{Q}_n$ captures approximately the distribution of the permuted data.[7] This consistency-type result requires that the information available from

[5]Condition (ii) is satisfied if $U_{n,i}$ has a conditional probability density that is uniformly bounded in probability.

[6]In our applications, we can often verify condition (iii) with $\widetilde{Y}_{n,i} = Y_{n,i}$. Nonetheless, allowing $\widetilde{Y}_{n,i} \neq Y_{n,i}$ is useful when $Y_{n,i}$ is itself an estimator. For example, if $(Y_{n,i})_{i\in\mathcal{I}_n}$ is a finite collection of estimators that converge jointly in distribution, then the coupling required in Assumption 2.1(iii) can be obtained via the Skorokhod representation. As such, the theory developed in Canay and Kamat (2017) can be cast in our framework with $\mathcal{G}_n$ being the trivial information set and k_n being a fixed constant, although our anticoncentration condition in Assumption 2.1(ii) is slightly stronger than the continuity condition in Canay and Kamat (2017, Assumption 4.2).

[7]When the two subsamples have different sizes, the same result goes through with $\overline{Q}_n$ defined as the sample-size weighted average between $Q_{1,n}$ and $Q_{2,n}$, given by $\overline{Q}_n = (Q_{1,n}|\mathcal{I}_{1,n}| + Q_{2,n}|\mathcal{I}_{2,n}|)/(|\mathcal{I}_{1,n}| + |\mathcal{I}_{2,n}|)$.

each subsample grows with the sample size, that is, $k_n \to \infty$. This result appears to be new in the context of permutation-based tests under a fixed alternative for the coupling variables. In particular, we note that an analogous result is unavailable in Canay and Kamat (2017), as they restrict attention to an asymptotic framework with a fixed k_n, which makes a consistency-type result unavailable. Our proof relies on applying Lehmann and Romano (2005, Theorem 15.2.3) to the infeasible test, for which we use the coupling construction developed by Chung and Romano (2013) to show that the so-called Hoeffding (1952) condition is satisfied. We note that this argument is used to establish the consistency of the permutation test rather than its asymptotic size property.

Theorem 2.1 establishes the relation between the rejection probability of the feasible test $\hat{\phi}_n$ and the homogeneity (or the lack of it) across the two coupling subsamples $(U_{n,i})_{i\in\mathcal{I}_{1,n}}$ and $(U_{n,i})_{i\in\mathcal{I}_{2,n}}$. This result does not speak directly to hypotheses formulated in terms of the original $(Y_{n,i})_{i\in\mathcal{I}_n}$ observations. Rather, its theoretical significance is to "absorb" all generic technicalities stemming from the (feasible) permutation test, which in turn considerably simplifies our overall analysis. The residual issue for any specific application is to explicitly construct the coupling variables and translate their homogeneity in terms of the primitive structures of the original empirical problem, which can be done using domain-specific techniques. We provide general results along this line in the infill time-series context, as detailed in Section 2.2 below.

To help anticipate the general discussion, it is instructive to sketch the scheme in a basic running example. Let $Y_{n,i} = \Delta_n^{-1/2}(P_{(i+1)\Delta_n} - P_{i\Delta_n})$ be the scaled increment of the asset price process P_t over the ith sampling interval. Let τ be a "cutoff" time point of interest that is known to the researcher (e.g., the announcement time of a news release) and i^* be the unique integer such that $\tau \in [i^*\Delta_n, (i^*+1)\Delta_n)$.[8] We consider two index sets $\mathcal{I}_{1,n} = \{i^* - k_n, \ldots, i^* - 1\}$ and $\mathcal{I}_{2,n} = \{i^* + 1, \ldots, i^* + k_n\}$, which collect observations before and after the cutoff, respectively. We consider an asymptotic setting in which these subsamples are "local" in calendar time, that is, $k_n\Delta_n \to 0$. Note that this implies that $\Delta_n \to 0$, which means that we are considering an infill asymptotic setting. If P_t is an Itô process with respect to an information filtration $(\mathcal{F}_t)_{t\geq 0}$, we may represent $Y_{n,i}$ as

$$Y_{n,i} = \Delta_n^{-1/2} \int_{i\Delta_n}^{(i+1)\Delta_n} b_s\, ds + \Delta_n^{-1/2} \int_{i\Delta_n}^{(i+1)\Delta_n} \sigma_s\, dW_s, \quad \text{for } i \in \mathcal{I}_n, \tag{2}$$

where b_t is the drift process, σ_t is the stochastic volatility process, and W_t is a standard Brownian motion.[9] If the σ_t process is smooth (e.g., Hölder continuous) in a local neighborhood before τ, then the volatility throughout the preevent subsample $\mathcal{I}_{1,n}$ is approximately $\sigma_{(i^*-k_n)\Delta_n}$. Further recognizing that the drift term is negligible relative to the Brownian component, we can approximate $Y_{n,i}$ for each $i \in \mathcal{I}_{1,n}$ using the coupling variables

$$U_{n,i} = \sigma_{(i^*-k_n)\Delta_n}\Delta_n^{-1/2}(W_{(i+1)\Delta_n} - W_{i\Delta_n}) \sim \mathcal{MN}\big(0,\, \sigma^2_{(i^*-k_n)\Delta_n}\big), \tag{3}$$

[8]The integer i^* depends on n. We suppress this dependence in our notation for simplicity and to avoid having nested subscripts.

[9]Note that $\mathcal{I}_n$ does not include the i^*th return observation. Therefore, although the returns in (2) do not contain price jumps, an event-induced price jump is allowed to occur at time τ.

where $\mathcal{MN}$ denotes the mixed normal distribution. Since the Brownian motion has independent and stationary increments, it is easy to see that the coupling variables $(U_{n,i})_{i\in\mathcal{I}_{1,n}}$ are $\mathcal{F}_{(i^*-k_n)\Delta_n}$-conditionally i.i.d. Moreover, if the volatility process σ_t does not jump at the cutoff time τ, we may follow the same logic to extend the approximation in (3) further to $i \in \mathcal{I}_{2,n}$. In other words, if the volatility process process does not jump then the coupling variables $(U_{n,i})_{i\in\mathcal{I}_n}$ are conditionally i.i.d., which corresponds to the situation in part (a) of Theorem 2.1. On the other hand, if the volatility process jumps at time τ, say by a constant $c \neq 0$, then the coupling variables for the $\mathcal{I}_{2,n}$ subsample will instead take the form $U_{n,i} = (\sigma_{(i^*-k_n)\Delta_n} + c)(W_{(i+1)\Delta_n} - W_{i\Delta_n})$. In this case, the two subsamples of $U_{n,i}$'s have distinct conditional distributions (i.e., mixed normal with different conditional variances), corresponding to the scenario in part (b) of Theorem 2.1.

Within the context of this illustrative example, we can further clarify a key feature of the proposed test that holds more generally. It is not aimed at detecting "small" time variations in the distribution of the observed data. In fact, by allowing the drift b_t and the volatility σ_t to be time varying, a smooth form of heterogeneity is *always* built in. The test instead detects abrupt changes, or discontinuities, in the evolution of the distribution, which can be more plausibly associated with the "lumpy" information carried by the underlying economic announcement, as emphasized by Nakamura and Steinsson (2018b). Specifically in this example, the asset returns are locally centered Gaussian (due to the assumption that the price is an Itô process), and hence, the temporal discontinuity in the return distribution manifests itself as a volatility jump. The empirical scope of our permutation test, however, is far beyond volatility-jump testing depicted in this illustration, as we shall demonstrate in the remainder of the paper.

2.2 *Permutation tests for discontinuities in event studies*

We now specialize the general Theorem 2.1 into an infill asymptotic time-series setting that is particularly suitable for event studies. By introducing a mild additional econometric structure, we shall establish the asymptotic validity of the permutation test under more primitive conditions that are easy to verify in a variety of concrete empirical settings. As in the running example above, we consider an event occurring at time $\tau \in [i^*\Delta_n, (i^*+1)\Delta_n)$, which separates two subsamples indexed by $\mathcal{I}_{1,n} = \{i^* - k_n, \ldots, i^* - 1\}$ and $\mathcal{I}_{2,n} = \{i^* + 1, \ldots, i^* + k_n\}$, respectively. All limits in the sequel are obtained under the infill asymptotic setting with $\Delta_n \to 0$.

We suppose that the data are generated from an approximate state-space model of the form

$$Y_{n,i} = g(\zeta_{i\Delta_n}, \epsilon_{n,i}) + R_{n,i}, \quad i \in \mathcal{I}_n, \tag{4}$$

where the state process ζ_t is càdlàg, adapted to a filtration $\mathcal{F}_t$, and takes values in an open set $\mathcal{Z} \subseteq \mathbb{R}^{\dim(\zeta)}$; $(\epsilon_{n,i})_{i\in\mathcal{I}_n}$ are i.i.d. random disturbances taking values in some (possibly abstract) space $\mathcal{E}$; $g(\cdot,\cdot)$ is a "smooth" transform; and $R_{n,i}$ is a residual term that is negligible relative to the leading term $g(\zeta_{i\Delta_n}, \epsilon_{n,i})$ in a proper sense detailed below. A simpler version of this state-space model without the $R_{n,i}$ residual term has been used by Li and Xiu (2016) and Bollerslev, Li, and Xue (2018), among others, for modeling market variables such as trading volume and bid-ask spread. By introducing the $R_{n,i}$ residual term, we can use a unified framework to accommodate a broader class of models, which

in particular include increments of an Itô semimartingale. We now revisit the model in (2) as the first illustration.

EXAMPLE 1 (Brownian Asset Returns). We represent the Itô process model (2) for asset returns in the form of (4) by setting $\zeta_t = \sigma_t$, $\epsilon_{n,i} = \Delta_n^{-1/2}(W_{(i+1)\Delta_n} - W_{i\Delta_n})$, and $g(z, \epsilon) = z\epsilon$. The resulting residual term has the form

$$R_{n,i} = \Delta_n^{-1/2} \int_{i\Delta_n}^{(i+1)\Delta_n} b_s \, ds + \Delta_n^{-1/2} \int_{i\Delta_n}^{(i+1)\Delta_n} (\sigma_s - \sigma_{i\Delta_n}) \, dW_s, \quad i \in \mathcal{I}_n. \tag{5}$$

Under mild and fairly standard regularity conditions, it is easy to show that $\max_{i\in\mathcal{I}_n} |R_{n,i}|$ is $o_p(1)$. On the other hand, the leading term $g(\zeta_{i\Delta_n}, \epsilon_{n,i})$ has a nondegenerate centered mixed Gaussian distribution with conditional variance $\sigma^2_{i\Delta_n}$.

This running example further illustrates the distinct roles played by ζ_t, $\epsilon_{n,i}$, and $R_{n,i}$ in our state-space model (4). The leading term $g(\zeta_{i\Delta_n}, \epsilon_{n,i})$ captures the "main feature" of the observed data; in addition, since the $\epsilon_{n,i}$ disturbance terms are i.i.d., any "large" change in the empirical distribution across the two subsamples must be attributed to the time-τ discontinuity in the state process ζ_t. From this description, it follows that the hypothesis test for the continuity of the distribution of the main feature of the observed data can be formulated as

$$H_0 : \Delta\zeta_\tau = 0 \quad \text{versus} \quad H_a : \Delta\zeta_\tau \neq 0, \tag{6}$$

where $\Delta\zeta_\tau \equiv \zeta_\tau - \zeta_{\tau-} \equiv \zeta_\tau - \lim_{s\uparrow\tau} \zeta_s$ denotes the jump of the state process at time τ.

With the state-space model (4) in place, we can design more primitive sufficient conditions for establishing the asymptotic validity of the permutation test under the hypotheses in (6). We need some additional notation to describe these conditions. For each fixed $z \in \mathcal{Z}$, let $f_z(\cdot)$ and $F_z(\cdot)$ denote the probability density function (PDF) and the CDF of the random variable $g(z, \varepsilon_{n,i})$, respectively. It is also convenient to introduce a "shifted" version of ζ_t defined as $\tilde{\zeta}_t \equiv \zeta_t - \Delta\zeta_\tau 1\{t \geq \tau\}$, which has the same increments as ζ_t over time intervals not containing τ.

ASSUMPTION 2.2. (*i*) *The collection of variables* $(\epsilon_{n,i})_{i\in\mathcal{I}_n}$ *are i.i.d. and, for each* $k \in \mathcal{I}_n$, *the variables* $(\epsilon_{n,i})_{i\geq k}$ *are independent of* $\mathcal{F}_{k\Delta_n}$. *Moreover, for any compact subset* $\mathcal{K} \subseteq \mathcal{Z}$, *we have* (*ii*) $\sup_{x\in\mathbb{R}, z\in\mathcal{K}} f_z(x) < \infty$ *and* (*iii*) $\inf_{z\in\mathcal{K}} \int_{\mathbb{R}} (F_z(x) - F_{z+c}(x))^2 \, dF_z(x) > 0$ *whenever* $c \neq 0$.

ASSUMPTION 2.3. *There exist a sequence* $(T_m)_{m\geq 1}$ *of stopping times increasing to infinity, a sequence of compact subsets* $(\mathcal{K}_m)_{m\geq 1}$ *of* $\mathcal{Z}$, *and a sequence* $(K_m)_{m\geq 1}$ *of constants such that for some real sequence* $a_n \geq 1$ *and each* $m \geq 1$: (*i*) $\|g(z, \epsilon_{n,i}) - g(z', \epsilon_{n,i})\|_2 \leq K_m a_n \|z - z'\|$ *for all* $z, z' \in \mathcal{K}_m$; (*ii*) ζ_t *takes values in* $\mathcal{K}_m$ *for all* $t \leq T_m$, *and* $\|\tilde{\zeta}_{t\wedge T_m} - \tilde{\zeta}_{s\wedge T_m}\|_2 \leq K_m |t - s|^{1/2}$ *for all* t, s *in some fixed neighborhood of* τ; (*iii*) $\max_{i\in\mathcal{I}_n} |R_{n,i}| = o_p(k_n^{-2})$.

Assumption 2.2 entails regularity conditions pertaining to the random disturbance terms, which are often easy to verify in concrete examples as demonstrated later in this subsection. Assumption 2.3 imposes a set of smoothness conditions that permits the

approximation of the observed data using properly constructed coupling variables.[10] Specifically, condition (i) requires that the random function $z \mapsto g(z, \epsilon_{n,i})$ is Lipschitz in z over compact sets under the L_2 distance. The a_n sequence captures the scale of the Lipschitz coefficient. In many applications, we can verify this condition simply with $a_n \equiv 1$, but allowing a_n to diverge to infinity is sometimes necessary (see Example 2 below). Condition (ii) states that the ζ_t process is locally compact (up to each stopping time T_m) and, upon removing the fixed-time discontinuity at τ, it is $(1/2)$-Hölder continuous under the L_2 norm. This Hölder-continuity requirement can be easily verified using well-known results provided that the $\tilde{\zeta}$ process is an Itô semimartingle or a long-memory process (see Jacod and Protter (2012, Chapter 2) and Li and Liu (2020)). Condition (iii) imposes the requisite assumptions on the residual terms. In some applications, this condition holds trivially with $R_{n,i} \equiv 0$, but, more generally, it needs to be verified on a case-by-case basis using (relatively standard) infill asymptotic techniques. Theorem 2.2, below, establishes the size and power properties of the permutation test under the hypotheses described in (6).

THEOREM 2.2. *In the state-space model* (4)*, suppose that Assumptions* 2.2 *and* 2.3 *hold, and that* $a_n k_n^3 \Delta_n^{1/2} = o(1)$*. Then the following statements hold for the permutation test* $\hat{\phi}_n$ *described in Algorithm* 1:

(*a*) *Under the null hypothesis in* (6)*, that is,* $\Delta\zeta_\tau = 0$*, we have* $\mathbb{E}[\hat{\phi}_n] \to \alpha$;

(*b*) *Under a fixed alternative hypothesis in* (6)*, that is,* $\Delta\zeta_\tau = c$ *for some* (*unknown*) *constant* $c \neq 0$ *and if* $k_n \to \infty$*, we have* $\mathbb{E}[\hat{\phi}_n] \to 1$.

This theorem is proved by verifying the high-level conditions in Theorem 2.1 with properly constructed coupling variables analogous to those in equation (3). The condition $a_n k_n^3 \Delta_n^{1/2} = o(1)$ mainly requires that the window size k_n does not grow too fast, which ensures the closeness between the coupling variables and the original data. In the typical case with $a_n = 1$, it reduces to $k_n = o(\Delta_n^{-1/6})$.[11] In general, a larger k_n allows one to utilize more data, but the associated longer event window may also lead to a larger nonparametric bias, and hence, a more severe size distortion. Part (a) shows that the permutation test attains the desired asymptotic level under the null hypothesis in (6). Again, we stress that the test has valid asymptotic size control even in the "small-sample" case with fixed k_n. As in Theorem 2.1, the "large-sample" condition $k_n \to \infty$ is needed for establishing the consistency of the test under the alternative, as shown in part (b).

In the remainder of this subsection, we use a few prototype examples to demonstrate how the proposed test may be used in various empirical settings. In particular, we show how to cast the specific problems into the approximate state-space model (4), and discuss how to verify our sufficient regularity conditions. We start by revisiting the running example.

[10]Note that the assumption is framed in a localized fashion using the stopping times $(T_m)_{m\geq 1}$, which is a standard technique for weakening the regularity condition in the infill asymptotic setting. See Jacod and Protter (2012, Section 4.4.1) for a comprehensive discussion on the localization technique.

[11]This sufficient condition for the growth rate of k_n is different from the conditions needed for conventional asymptotic-Gaussian-based spot inference, which requires $k_n \asymp \Delta_n^{-\iota}$ for some $\iota \in (0, 1/2)$. For the permutation test, k_n may be fixed or grow to infinity. However, in the latter case, our condition on the growth rate of k_n is more stringent than what is needed for the conventional spot inference theory.

EXAMPLE 1 (Brownian Asset Returns, Continued). Recall that $\epsilon_{n,i} \equiv \Delta_n^{-1/2}(W_{(i+1)\Delta_n} - W_{i\Delta_n})$, $\zeta_t = \sigma_t$, and $g(z, \epsilon) = z\epsilon$. In this context, the hypothesis testing problem in (6) represents a test of the continuity of the volatility process σ_t at time $t = \tau$, that is,

$$H_0 : \Delta\sigma_\tau = 0 \quad \text{versus} \quad H_a : \Delta\sigma_\tau \neq 0.$$

We suppose that the volatility process σ_t is nondegenerate by setting its domain to $\mathcal{Z} = (0, \infty)$. Since the Brownian motion has independent increments with respect to the underlying filtration, the disturbance term $\epsilon_{n,i}$ satisfies Assumption 2.2(i). In addition, for each point $z \in \mathcal{Z}$, the random variable $f(z, \epsilon_{n,i})$ has an $\mathcal{N}(0, z^2)$ distribution. It is then easy to see that conditions (ii) and (iii) in Assumption 2.2 hold for any compact subset $\mathcal{K} \subseteq \mathcal{Z}$ (note that $\mathcal{K}$ is necessarily bounded away from zero). To verify Assumption 2.3, first note that $g(z, \epsilon_{n,i}) - g(z', \epsilon_{n,i}) = (z - z')\epsilon_{n,i}$, and hence, $\|g(z, \epsilon_{n,i}) - g(z', \epsilon_{n,i})\|_2 = |z - z'|$. Assumption 2.3(i) thus holds for $a_n = 1$. It is well known that σ_t is locally $(1/2)$-Hölder continuous under the L_2 norm if it is an Itô semimartingale or a long-memory process; if so, Assumption 2.3(ii) is satisfied if the σ_t and σ_t^{-1} processes are both locally bounded. Finally, to verify Assumption 2.3(iii), we assume that the drift process b_t is locally bounded. It is then easy to show via routine calculations that $\max_{i\in\mathcal{I}_n} |R_{n,i}| = O_p(k_n^{1/2}\Delta_n^{1/2})$. Since the condition $a_n k_n^3 \Delta_n^{1/2} = o(1)$ in Theorem 2.2 implies that $O_p(k_n^{1/2}\Delta_n^{1/2}) = o_p(k_n^{-2})$, we have $\max_{i\in\mathcal{I}_n} |R_{n,i}| = o_p(k_n^{-2})$ as needed in Assumption 2.3(iii). All conditions in Theorem 2.2 are now verified, and this shows that the permutation test $\hat{\phi}_n$ is asymptotically valid for testing the null hypothesis $\Delta\sigma_\tau = 0$.

Example 1 shows that the permutation test $\hat{\phi}_n$ is asymptotically valid for testing the presence of a volatility jump. This is a relatively familiar problem in the literature. It is therefore useful to contrast the proposed permutation test with the standard approach, which is based on nonparametric "spot" estimators of the asset price's instantaneous variances before and after the event time given by, respectively,

$$\hat{\sigma}_{\tau-}^2 = \frac{1}{k_n} \sum_{i\in\mathcal{I}_{1,n}} Y_{n,i}^2, \qquad \hat{\sigma}_{\tau}^2 = \frac{1}{k_n} \sum_{i\in\mathcal{I}_{2,n}} Y_{n,i}^2. \tag{7}$$

Assuming $k_n \to \infty$ and $k_n^2\Delta_n \to 0$, it can be shown that (see Jacod and Protter (2012, Chapter 13))

$$\frac{k_n^{1/2}\big(\hat{\sigma}_\tau^2 - \hat{\sigma}_{\tau-}^2 - \big(\sigma_\tau^2 - \sigma_{\tau-}^2\big)\big)}{\sqrt{2\hat{\sigma}_\tau^4 + 2\hat{\sigma}_{\tau-}^4}} \xrightarrow{d} \mathcal{N}(0, 1). \tag{8}$$

Thus, we can test $H_0 : \Delta\sigma_\tau = 0$ by comparing the t-statistic $k_n^{1/2}(\hat{\sigma}_\tau^2 - \hat{\sigma}_{\tau-}^2)/\sqrt{2\hat{\sigma}_\tau^4 + 2\hat{\sigma}_{\tau-}^4}$ with critical values based on the standard normal distribution.

Two remarks are in order. First, note that the asymptotic size control of the standard approach relies on the asymptotic normal approximation (8), which depends crucially on $k_n \to \infty$ (in addition to having $\Delta_n \to 0$) because the underlying central limit theorem is obtained by aggregating a "large" number of martingale differences. Hence, the t-test may suffer from severe size distortion when k_n is relatively small. This issue is empir-

ically relevant because an applied researcher may use a short time window to capture short-lived "impulse-like" dynamics and/or to minimize the impact of other confounding economic factors in the background. Moreover, for "real-time" applications, the researcher may have no choice but to use a small k_n simply because of the limited amount of available data soon after the event time τ. In sharp contrast, the permutation test controls asymptotic size even when k_n is fixed. This remarkable property is inherited from the coupling two-sample problem, in which the permutation test controls size exactly regardless of whether k_n is fixed or grows to infinity.

The second and perhaps practically more important difference between the two tests is that the permutation test is more versatile. Under the spot-estimation-based approach, both the design of the spot estimators in (7) and the convergence in (8) depend heavily on the fact that the increments of the Brownian motion are not only i.i.d., but also Gaussian. Gaussianity is obviously essential for the conventional approach because, among other things, it ensures that the instantaneous variance of the normalized returns are well-defined.[12] The permutation test, on the other hand, only exploits the i.i.d. property of the Brownian shocks, without relying on their Gaussianity. Therefore, the permutation test readily accommodates a more general model for asset returns with Lévy shocks, as we demonstrate in the following example.

EXAMPLE 2 (Lévy-driven Asset Returns). We generalize the model in Example 1 by replacing the Brownian motion W with a Lévy martingale L, so that the asset return has the form

$$P_{(i+1)\Delta_n} - P_{i\Delta_n} = \int_{i\Delta_n}^{(i+1)\Delta_n} b_s \, ds + \int_{i\Delta_n}^{(i+1)\Delta_n} \sigma_s \, dL_s, \quad i \in \mathcal{I}_n.$$

In this case, we define the random disturbance as $\epsilon_{n,i} \equiv \Delta_n^{-1/\beta}(L_{(i+1)\Delta_n} - L_{i\Delta_n})$ for some constant $\beta \in (1, 2]$. The more general normalizing sequence $\Delta_n^{-1/\beta}$ is used to ensure that $\epsilon_{n,i}$ has a non-degenerate distribution. For instance, if L is a stable process, we take β to be its jump-activity index, so that $\epsilon_{n,i}$ has a centered stable distribution (recall that the Brownian motion is a stable process with index $\beta = 2$). We treat the value of β as unknown. Since the permutation test is scale-invariant with respect to the data, we can nonetheless regard the normalized return $Y_{n,i} = \Delta_n^{-1/\beta}(P_{(i+1)\Delta_n} - P_{i\Delta_n})$ as directly observable (because tests implemented for $P_{(i+1)\Delta_n} - P_{i\Delta_n}$ and $Y_{n,i}$ are identical). To apply our theory, we represent $Y_{n,i}$ using the state-space model (4) with $\zeta_t = \sigma_t$, $g(z, \epsilon) = z\epsilon$, and the residual term given by

$$R_{n,i} = \Delta_n^{-1/\beta} \int_{i\Delta_n}^{(i+1)\Delta_n} b_s \, ds + \Delta_n^{-1/\beta} \int_{i\Delta_n}^{(i+1)\Delta_n} (\sigma_s - \sigma_{i\Delta_n}) \, dL_s, \quad i \in \mathcal{I}_n.$$

Recognizing that the scaled Lévy increments $(\epsilon_{n,i})_{i\in\mathcal{I}_n}$ are i.i.d., we can verify Assumptions 2.2 and 2.3 using similar arguments as in Example 1 but with $a_n = \Delta_n^{1/2-1/\beta}$, which

[12] Recall that many distributions used in continuous-time models do not have finite second moments. For example, within the class of stable distributions, the Gaussian distribution is the only one with a finite second moment. Moreover, Gaussianity also implies that the variance of $\Delta_n^{-1}(W_{i\Delta_n} - W_{(i-1)\Delta_n})^2$ is 2, which explains the "2" factor in the denominator of the t-statistic.

depicts the rate at which $\|\epsilon_{n,i}\|_2$ diverges. In particular, the condition $a_n k_n^3 \Delta_n^{1/2} = o(1)$ requires k_n to obey $k_n = o(\Delta_n^{(1/\beta-1)/3})$. Then we can apply Theorem 2.2 to show that the permutation test $\hat{\phi}_n$ is asymptotically valid for testing the discontinuity in the volatility process σ_t at time τ, regardless of whether the driving Lévy process is a Brownian motion or not. To our knowledge, our test is the first in the literature that accommodates Lévy-type shocks in the context of testing for volatility jumps.

So far, we have illustrated the use of the permutation test for high-frequency asset returns data. Under the settings of Examples 1 and 2, the distributional change of asset returns is mainly driven by the time-τ discontinuity in volatility, and hence, the permutation test is effectively a test for volatility jumps. Example 2, in particular, highlights the versatility and robustness of the permutation test compared with the conventional approach based on spot estimation. Going one step further, we now illustrate how to apply the permutation test to other types of economic variables.

Example 3 (Location-Scale Model for Volume). Consider a simple model for trading volume, under which the volume within the ith sampling interval is given by $Y_{n,i} = \mu_{i\Delta_n} + v_{i\Delta_n}\epsilon_{n,i}$. The μ_t location process captures the local mean, or trading intensity, and the v_t scale process captures the time-varying heterogeneity in the order size. This location-scale model fits directly into the state-space model (4) with $\zeta_t = (\mu_t, v_t)$, $g((\mu, v), \epsilon) = \mu + v\epsilon$, and $R_{n,i} \equiv 0$. Let $\mathcal{F}_t$ be the filtration generated by the ζ_t process. If $\epsilon_{n,i}$ is independent of the ζ_t process and has finite second moment and bounded PDF, then it is easy to verify Assumptions 2.2 and 2.3 with $a_n = 1$. Theorem 2.2 thus implies that the permutation test is valid for testing the discontinuity in $\zeta_t = (\mu_t, v_t)$ at time τ.

The location-scale structure in Example 3 is by no means essential in applications, because the permutation test is valid provided that the more general conditions in Assumptions 2.2 and 2.3 hold. This illustration is pedagogically convenient, in that it permits a straightforward verification of our high-level conditions. That being said, this example does reveal a limitation of our theory developed so far. That is, the data variable needs to be continuously distributed, as required in Assumption 2.2(ii) (which in turn is related to Assumption 2.1(ii)). Observed data in actual applications are invariably discrete, but this continuous-distribution assumption is often deemed as a reasonable approximation to reality. In some situations, however, the discreteness in the data is more salient. For example, the trading volume of a relatively illiquid asset may take values as small integer multiples of the lot size (e.g., 100 shares).[13] This motivates us to directly confront the discreteness in the data, as detailed in the next subsection.

2.3 *Extension: The case with discretely valued data*

The extension will be carried out in similar steps as the theory developed above. We start with modifying the general result in Theorem 2.1 to accommodate discretely valued

[13]This issue has become less important in the equity market as retail investors can now trade a single share, or even a fractional share, of a stock. However, the lot size is still relevant for less liquid assets such as option contracts or for equity data from earlier sample periods.

observations; we then specialize the general theory to a high-frequency setting under more primitive conditions. Recall that $Q_{j,n}(\cdot)$ denotes the $\mathcal{G}_n$-conditional distribution function of the coupling variable $U_{n,i}$ for $i \in \mathcal{I}_{j,n}$ and $j \in \{1, 2\}$, and $\overline{Q}_n = (Q_{1,n} + Q_{2,n})/2$.

THEOREM 2.3. *Suppose that there exists a collection of variables $(U_{n,i})_{i\in\mathcal{I}_n}$ that satisfies Assumption 2.1(i) for some sequence $(\mathcal{G}_n)_{n\geq 1}$ of σ-fields, and $\mathbb{P}(\widetilde{Y}_{n,i} \neq U_{n,i}) = o(k_n^{-1})$ uniformly in $i \in \mathcal{I}_n$ where $(\widetilde{Y}_{n,i})_{i\in\mathcal{I}_n}$ is an identical copy of $(Y_{n,i})_{i\in\mathcal{I}_n}$ in $\mathcal{G}_n$-conditional distribution. Then the following statements hold for the test $\hat{\phi}_n$ described in Algorithm 1:*

(a) If the variables $(U_{n,i})_{i\in\mathcal{I}_n}$ have the same $\mathcal{G}_n$-conditional distribution, we have $\mathbb{E}[\hat{\phi}_n] \to \alpha$.

(b) If $k_n \to \infty$ and $\mathbb{P}(\int (Q_{1,n}(x) - Q_{2,n}(x))^2 \, d\overline{Q}_n(x) > \delta_n) \to 1$ for any real sequence $\delta_n = o(1)$, we have $\mathbb{E}[\hat{\phi}_n] \to 1$.

Theorem 2.3 establishes exactly the same asymptotic properties for the permutation test as Theorem 2.1, but under different conditions: it does not impose the anticoncentration requirement for the coupling variable (i.e., Assumption 2.1(ii)), and the "distance" between the observed data and the coupling variable is measured by the probability mass of $\{\widetilde{Y}_{n,i} \neq U_{n,i}\}$. These modifications seem natural for the discrete-data setting.

Next, we specialize the general result in Theorem 2.3 to the state-space model (4), starting with some motivating examples. The first is an alternative model for the trading volume that explicitly features discretely valued data, which shows an interesting contrast to Example 3.

EXAMPLE 4 (Poisson Model for Volume). Let $Y_{n,i}$ be the trading volume of an asset within the ith sampling interval. Following Andersen (1996), we model the discretely valued volume using a Poisson distribution with time-varying mean. To form a state-space representation, let $(\epsilon_{n,i}(t))_{t\geq 0}$ be a copy of the standard Poisson process on $\mathbb{R}_+$, independent across i, and let ζ_t be the time-varying mean process independent of the $\epsilon_{n,i}$'s. We then set $Y_{n,i} = \epsilon_{n,i}(\zeta_{i\Delta_n})$, which conditional on the ζ process, is Poisson distributed with mean $\zeta_{i\Delta_n}$. This representation is a special case of (4), with $g(\zeta, \epsilon) = \epsilon(\zeta)$ being a time change and $R_{n,i} = 0$. We also note that although the $\epsilon_{n,i}$'s are assumed to be i.i.d., the $(Y_{n,i})_{i\in\mathcal{I}_n}$ series can be highly persistent through its dependence on the stochastic mean process ζ_t.

To further broaden the empirical scope, we consider another example concerning the bid-ask spread of asset quotes. This example is econometrically interesting because of its resemblance to the discrete-choice models (e.g., probit and logit) commonly used for modeling binary and multinomial data.

EXAMPLE 5 (Bid-Ask Spread). Let $Y_{n,i}$ be the bid-ask spread of an asset at time $i\Delta_n$. For a liquid asset, the spread is often maintained at 1 tick (e.g., 1 cent), but it may widen to several ticks due to a higher level of asymmetric information or dealer's inventory cost. For ease of exposition, we suppose that $Y_{n,i}$ is a binary variable taking values in $\{1, 2\}$,

while noting that a multinomial extension is straightforward. Motivated by the classical discrete-choice models, we model the spread as $Y_{n,i} = 1 + 1\{\zeta_{i\Delta_n} \geq \epsilon_{n,i}\}$, and suppose that the variables $(\epsilon_{n,i})_{i\in\mathcal{I}_n}$ are i.i.d. and independent of the ζ_t process. With the CDF of $\epsilon_{n,i}$ denoted by $F_\epsilon(\cdot)$, we have $\mathbb{P}(Y_{n,i} = 2|\zeta_{i\Delta_n}) = F_\epsilon(\zeta_{i\Delta_n})$. Evidently, upon redefining ζ_t as $F_\epsilon(\zeta_t)$, we can assume that $\epsilon_{n,i}$ is uniformly distributed on the $[0, 1]$ interval without loss of generality. This normalization in turn allows us to interpret ζ_t as the stochastic propensity of a "wide" spread, which may serve as a measure of market illiquidity.

We now proceed to establish the asymptotic validity of the permutation test for the hypotheses described in (6) for discretely valued observations; see Theorem 2.4 below. Since the state-space representation (4) holds with the residual term $R_{n,i} = 0$ in the examples above, it seems reasonable to avoid unnecessary redundancy by restricting our analysis to a simpler version given by

$$Y_{n,i} = g(\zeta_{i\Delta_n}, \epsilon_{n,i}), \quad i \in \mathcal{I}_n. \tag{9}$$

We replace Assumption 2.3 with the following assumption, where we recall that for each $z \in \mathcal{Z}$, $F_z(\cdot)$ denotes the CDF of the random variable $g(z, \varepsilon_{n,i})$ and $\tilde{\zeta}_t = \zeta_t - \Delta\zeta_\tau 1\{t \geq \tau\}$.

Assumption 2.4. *There exist a sequence $(T_m)_{m\geq 1}$ of stopping times increasing to infinity, a sequence of compact subsets $(\mathcal{K}_m)_{m\geq 1}$ of $\mathcal{Z}$, and a sequence $(K_m)_{m\geq 1}$ of constants such that for each $m \geq 1$: (i) $\mathbb{P}(g(z, \epsilon_{n,i}) \neq g(z', \epsilon_{n,i})) \leq K_m\|z - z'\|$ for all $z, z' \in \mathcal{K}_m$; (ii) ζ_t takes values in $\mathcal{K}_m$ for all $t \leq T_m$, and $\|\tilde{\zeta}_{t\wedge T_m} - \tilde{\zeta}_{s\wedge T_m}\|_2 \leq K_m|t - s|^{1/2}$ for all t, s in some fixed neighborhood of τ.*

Theorem 2.4. *In the state-space model (9), suppose that Assumptions 2.2(i), 2.2(iii), and 2.4 hold, and that $k_n^3\Delta_n = o(1)$. Then the following statements hold for the permutation test $\hat{\phi}_n$ described in Algorithm 1:*

(a) Under the null hypothesis in (6), that is, $\Delta\zeta_\tau = 0$, we have $\mathbb{E}[\hat{\phi}_n] \to \alpha$;

(b) Under a fixed alternative hypothesis in (6), that is, $\Delta\zeta_\tau = c$ for some (unknown) constant $c \neq 0$, and if $k_n \to \infty$, we have that $\mathbb{E}[\hat{\phi}_n] \to 1$.

Theorem 2.4 depicts the same asymptotic behavior of the permutation test as in Theorem 2.2. The sufficient conditions of these results differ mainly in how to gauge the closeness between the data and the coupling variable, as manifest in the difference between Assumption 2.3(i) and Assumption 2.4(i). The latter is easy to verify under more primitive conditions in concrete settings. Specifically, in Example 4, we note that $|g(z, \epsilon_{n,i}) - g(z', \epsilon_{n,i})|$ is a Poisson random variable with mean $|z - z'|$, and hence, $\mathbb{P}(g(z, \epsilon_{n,i}) \neq g(z', \epsilon_{n,i})) = 1 - \exp(-|z - z'|) \leq |z - z'|$ as desired. In Example 5, we can use $\epsilon_{n,i} \sim \text{Uniform}[0, 1]$ to deduce that

$$\mathbb{P}\big(g(z, \epsilon_{n,i}) \neq g\big(z', \epsilon_{n,i}\big)\big) = \mathbb{P}\big(1\{z \geq \epsilon_{n,i}\} \neq 1\big\{z' \geq \epsilon_{n,i}\big\}\big) = \big|z - z'\big|,$$

which again verifies Assumption 2.4(i). Therefore, in the context of Examples 4 and 5 above, the permutation test is asymptotically valid for detecting discontinuities in trading activity and illiquidity, respectively.

3. Monte Carlo simulations

3.1 *Setting*

Our Monte Carlo experiment is based on the setting of Example 2. We simulate the (log) price process according to $dP_t = \sigma_t\, dL_t$ under an Euler scheme on a 1-second mesh, and then resample the data at the $\Delta_n = 1$ minute frequency. We simulate L either as a standard Brownian motion or as a (centered symmetric) stable process with index $\beta = 1.5$. To avoid an unrealistic price path, we truncate the stable distribution so that its normalized increment $\Delta_n^{-1/\beta}(L_{i\Delta_n} - L_{(i-1)\Delta_n})$ is supported on $[-C, C]$, and we consider $C \in \{10, 20, 30\}$ to examine the effect of the support. The unit of time is one day.

To simulate the volatility process, we first simulate two volatility factors according to the following dynamics (see Bollerslev (2011)):

$$dV_{1,t} = 0.0116(0.5 - V_{1,t})\, dt + 0.1023\sqrt{V_{1,t}}\big(\rho\, dL_t + \sqrt{1-\rho^2}\, dB_{1,t}\big) + c \cdot 1_{\{t=\tau\}},$$

$$dV_{2,t} = 0.6930(0.5 - V_{2,t})\, dt + 0.7909\sqrt{V_{2,t}}\big(\rho\, dL_t + \sqrt{1-\rho^2}\, dB_{2,t}\big) + c \cdot 1_{\{t=\tau\}},$$

where $B_{1,t}$ and $B_{2,t}$ are independent standard Brownian motions that are also independent of L_t, $\rho = -0.7$ captures the negative correlation between price and volatility shocks (namely, the "leverage" effect). The V_1 volatility factor is highly persistent with a half-life of 2.5 months, while the V_2 volatility factor is quickly mean-reverting with a half-life of only one day. The constant c determines the size of the volatility jump at the event time τ. In particular, $c = 0$ corresponds to the null hypothesis, and we consider a range of c values in $(0, 5]$ in order to trace out a power curve for the corresponding alternative hypotheses. The range of the c parameter is calibrated according to Bollerslev, Li, and Xue's (2018) empirical estimates for FOMC announcements.[14]

We note that the two volatility factors, V_1 and V_2, capture the slow- and fast-mean-reverting volatility dynamics, respectively, with the former having "smoother" sample paths than the latter. With this in mind, we simulate σ_t using two models:

$$\begin{cases} \text{Model A:} & \sigma_t^2 = 2V_{1,t}, \\ \text{Model B:} & \sigma_t^2 = V_{1,t} + V_{2,t}. \end{cases} \tag{10}$$

In finite samples, Model A features relatively smooth volatility paths, which is close to the "ideal" scenario underlying the infill asymptotic theory. Meanwhile, Model B generates more realistic, and rougher, sample path for σ, providing a nontrivial challenge for the proposed inference theory.

We implement the permutation test at the 5% significance level, with the window size $k_n \in \{15, 30, 60, 90\}$. The six-fold increase from the smallest window size to the largest one represents a considerable range that allows us to explore the robustness of

[14] Specifically, Bollerslev, Li, and Xue (2018) estimate the average jump size of $\log(\sigma_t)$ for the S&P 500 ETF around FOMC announcements to be 1.037 (see Table 3 of that paper). This suggests that $\sigma_\tau^2/\sigma_{\tau-}^2 = (\exp(1.037))^2 \approx 8$ on average, corresponding to $c \approx 3.5$ in this Monte Carlo design.

the proposed test with respect to the k_n tuning parameter.[15] The critical value is computed as in Remark 2.1 based on 1000 i.i.d. permutations. For comparison, we also implement the standard (two-sided) t-test based on (8). Rejection frequencies are computed based on 2000 Monte Carlo trials.

3.2 *Results*

We first examine the size properties of the permutation test $\hat{\phi}_n$ and the t-test based on (8). Table 1 reports the rejection frequencies of these tests under the null hypothesis (i.e., $c = 0$) for various data generating processes. Column (1) corresponds to the case with L being a standard Brownian motion, and columns (2), (3), and (4) report results when L is a truncated stable process with the truncation parameter $C = 10$, 20, and 30, respectively.

The top panel of the table shows results from Model A, where the volatility is solely driven by the "slow" factor. Quite remarkably, the rejection frequencies of the permutation test are very close to the 5% nominal level for all specifications of L, and importantly, for a wide range of the window size k_n. In contrast, the rejection rates of the t-test appear to be far more sensitive to the choice of k_n. As we increase k_n from 15 to 90, the rejection rate increases from 1.1% to 4.8% when L is a Brownian motion. A similar pattern emerges when L is a truncated stable process, except that the rejection rates now exceed the nominal level and reach 7.4% and 9.1% in columns (3) and (4), respectively. It is relevant to note that the t-test is not formally justified when L is not a Brownian motion.

TABLE 1. Rejection rates under the null hypothesis.

	Permutation test				T-test			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
				Model A: One-factor volatility				
$k_n = 15$	0.050	0.058	0.053	0.062	0.011	0.017	0.021	0.032
$k_n = 30$	0.054	0.049	0.056	0.057	0.031	0.042	0.044	0.063
$k_n = 60$	0.047	0.048	0.053	0.059	0.046	0.053	0.055	0.077
$k_n = 90$	0.058	0.050	0.053	0.056	0.048	0.056	0.074	0.091
				Model B: Two-factor volatility				
$k_n = 15$	0.049	0.055	0.054	0.062	0.014	0.017	0.023	0.037
$k_n = 30$	0.055	0.050	0.055	0.056	0.037	0.041	0.051	0.073
$k_n = 60$	0.049	0.049	0.054	0.064	0.091	0.077	0.078	0.102
$k_n = 90$	0.064	0.052	0.058	0.059	0.136	0.101	0.117	0.147

Note: This table presents rejection frequencies of the permutation test and the t-test under the null hypothesis $\sigma^2_{\tau-} = \sigma^2_{\tau}$. The significance level is fixed at 5%. Column (1) corresponds to the case with L being a standard Brownian motion, and columns (2)–(4) correspond to cases in which L is truncated stable with index 1.5 and truncation parameter $C \in \{10, 20, 30\}$. The rejection frequencies are computed based on 2000 Monte Carlo trials.

[15]We also implemented simulations in which the two subsamples, $\mathcal{I}_{1,n}$ and $\mathcal{I}_{2,n}$, have different sample sizes $k_{1,n}, k_{2,n} \in \{15, 30, 60, 90\}$. As anticipated in footnote 4, these results are quantitatively similar to those with a common sample size, that is, $k_{1,n} = k_{2,n}$. These additional results are omitted for brevity but are available upon request.

The more challenging case is Model B with the two-factor volatility dynamics. Looking at the bottom panel of Table 1, we find that the permutation test still has rejection rates that are quite close to the nominal level, although we see a slight overrejection of 6.4% when $k_n = 90$. This is likely due to the fact that the approximation error in the coupling has nontrivial impact when the window size is large. That being said, we note that the benchmark t-test is more severely affected by this bias issue, with rejection rates reaching 9.1% and 13.6% when $k_n = 60$ and $k_n = 90$, respectively.

Figure 1 plots the power curves of the permutation test and the t-test for various k_n's in Model A and Model B. We note that the four specifications of L produce qualitatively similar results. We see that the rejection frequencies increase with the window size k_n and the jump size c, which is expected from our consistency result obtained under $k_n \to \infty$. The permutation test appears to be less powerful than the t-test under the alternative hypothesis. This is a natural consequence of the typical trade-off between efficiency and robustness to the assumptions. The t-test is based on the spot variance estimator, which is "locally" the maximum-likelihood estimator of the spot variance under the Brownian shocks. The asymptotic validity and efficiency of the t-test rely on the

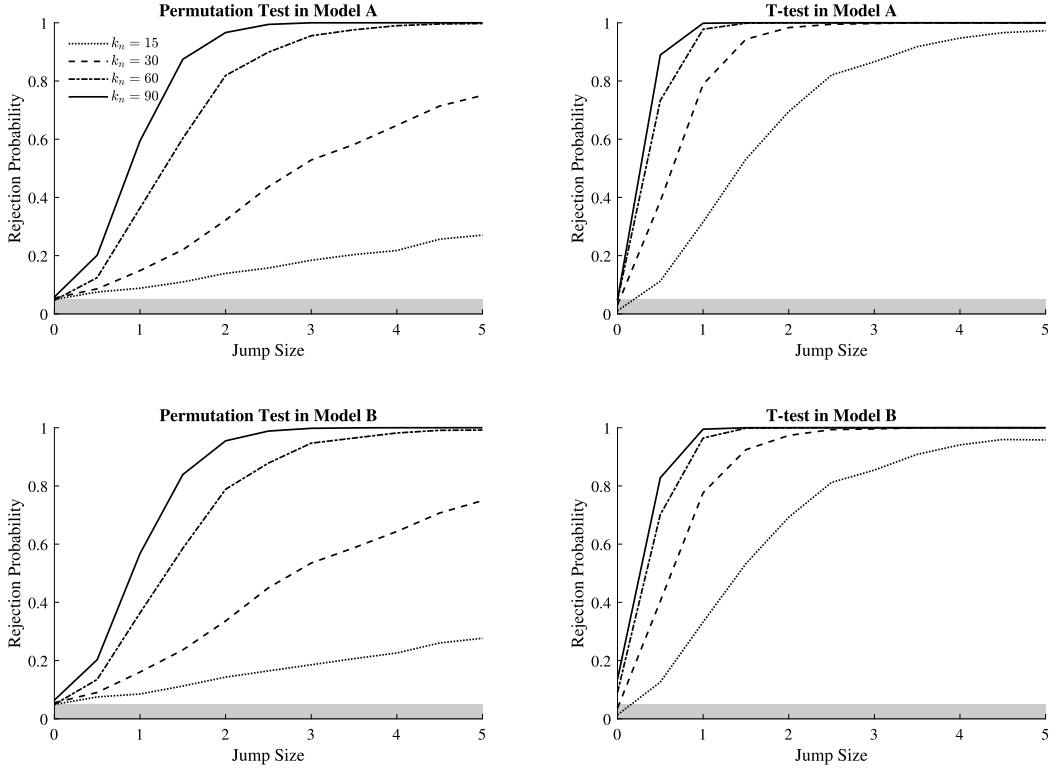

FIGURE 1. The figure plots the rejection frequencies of the permutation test and the t-test. The significance level is fixed at 5% (highlighted by shade). Results for Model A and Model B are presented in the top and bottom rows, respectively. In all the plots, L is simulated as a standard Brownian motion. The power curves are computed for the jump size parameter $c \in \{0, 0.5, 1, \ldots, 5\}$. The rejection frequencies are computed based on 2000 Monte Carlo trials.

Brownian assumption. In contrast, the permutation test is asymptotically valid regardless of whether the shocks are Brownian or not. As expected, this robustness is costly in terms of power. On the flip side, the power advantage of the t-test comes at the cost of size distortion when shocks are non-Brownian, which can be large, as shown in Table 1.

Overall, we find that the permutation test controls size remarkably well under the null hypothesis. Although it appears to be less powerful than the t-test, it does not suffer from the latter's size distortion, which can be severe in the two-factor volatility model. Our results suggest that, given its robustness, the permutation test is a useful complement to the conventional test based on spot estimation and asymptotic Gaussian approximation.

4. An empirical illustration

We apply the proposed permutation test to a recent sample of high-frequency price and volume observations of the S&P 500 ETF (NYSE: SPY) as an empirical illustration. The sampling frequency is 1 minute; the data source is the TAQ database. With the permutation test, we are interested in testing distributional discontinuities of the ETF's return, trading volume, and two measures of illiquidity for several FOMC announcements during the COVID-19 pandemic. This setting highlights one of the key merits of the proposed test, namely, it is applicable for a broad variety of high-frequency observations modeled in distinct ways. This is in sharp contrast to the conventional t-test based on (8), which is designed specifically for testing volatility jumps and whose validity relies on the assumption of Brownian shocks. We construct the high-frequency volume series as the total number of shares within each 1-minute trading session. The illiquidity measures of interest include Amihud's measure defined as the ratio between absolute return and dollar trading volume (Amihud (2002)) and the bid-ask spread averaged within each 1-minute trading session.

We focus on four important FOMC announcements in the 2020–2021 sample period that are related to distinct aspects of the Federal Reserve's monetary policy during the COVID-19 pandemic. The first is the announcement made on March 3, 2020, which was also the first FOMC announcement after COVID-19 hit the United States. The Fed stated its decision to lower the federal funds rate by 1/2 percentage point as its first response to counter the pandemic's negative impact on the economy. The second event occurred on December 16, 2020. At that time, being concerned with the rising long-term yield, many market participants expected that the Fed might implement the so-called "operation twist" to tame the steepening of the yield curve. But this turned out not to be on the Fed's agenda, and so the Fed's "inaction" may be deemed as a shock relative to the market's anticipation. The third case pertains to the announcement on March 17, 2021. During the press conference, the Fed Chairman suggested that the central bank would be unlikely to raise the rate in the next 2–3 years, which may be regarded as a forward guidance on the target rate. The final example is the announcement on September 22, 2021, when the Fed officially declared its intention to taper the large-scale asset purchase program.

Figure 2 plots the asset return and trading volume of SPY over 1-hour windows centered at these announcement times. For ease of comparison, we plot the return and

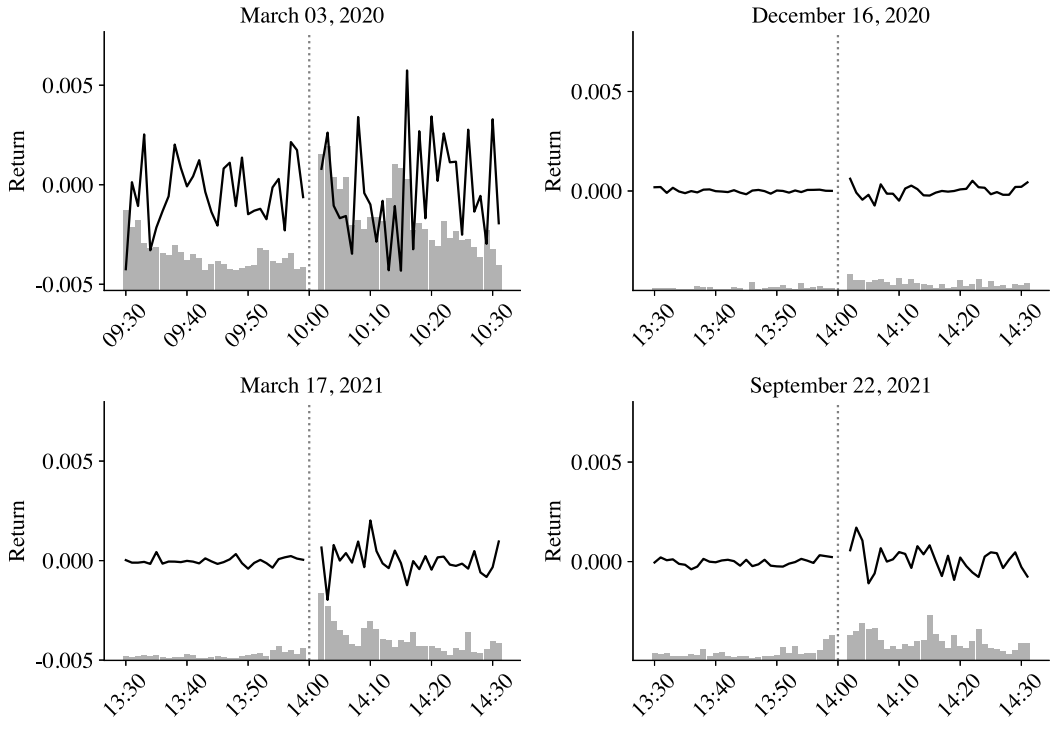

FIGURE 2. This figure plots the price return (solid) and volume (shaded) of the SPY ETF in 1-hour windows around four FOMC announcements.

volume data for the four events on the same scale. These plots immediately reveal the highly distinct market conditions at those times. This highlights the usefulness of adopting a high-frequency event-study research design, which allows us to investigate each event separately, rather than pooling information across different announcements under a likely fragile homogeneity assumption. We also observe several interesting patterns regarding how the market responds to the "lumpy" information embedded in the announcements. We generally see a rise in trading activity after the announcement. The price also tends to fluctuate more in the post-announcement window, although the March 3, 2020, event may be an exception as the market was already quite volatile even before the announcement.

As mentioned above, we implement the permutation test described in Algorithm 1 on the ETF's returns, trading volume, Amihud's measure, and bid-ask spread constructed on the 1-minute sampling frequency. For ease of interpretation, we consider the nonrandomized version of the test described in Remark 2.1. We consider two event windows: $k_n = 10$ minutes or 30 minutes. Recall that each FOMC announcement contains two stages. The first is an immediate release of a short summary on the Federal Reserve's webpage, which will be further detailed in the Fed Chairman's opening statement during the first (roughly) 10 minutes of the press conference. The second part is a Q&A session in which the Chairman responds to questions from the media, with the first few generally being more important. Given this setup, the shorter $k_n = 10$ window allows us

TABLE 2. Test results for selected FOMC announcements.

	$k_n = 10$				$k_n = 30$			
Date (τ)	Return	Volume	Amihud	Spread	Return	Volume	Amihud	Spread
03/03/2020		***	*	***		***		***
12/16/2020	**	***	**	***	**	***		***
03/17/2021	*	***		***	**	***		***
09/22/2021	**	***		***	***	***		***

Note: The permutation test is implemented following Algorithm 1. For ease of interpretation, we consider the nonrandomized version as described in Remark 2.1 using 100,000 i.i.d. permutations. Rejections at the 10%, 5%, and 1% significance levels are indicated by *, **, and ***, respectively.

to focus on the immediate impact of the FOMC statement, whereas the longer $k_n = 30$ window further covers the "more subtle" policy information conveyed to the public during the Q&A session. It is worth noting that the 30-minute event window is also adopted in prior work on the high-frequency identification of monetary policy shocks; see Nakamura and Steinsson (2018a).

Table 2 reports the results of the permutation test. Recall that the permutation test applied to asset returns may be interpreted as a test for volatility jumps, as explained in Example 1 (with Brownian shocks) or Example 2 (with more general Lévy shocks).[16] At a significance level of 10%, the permutation test rejects the null of continuity for all events except the one on March 3, 2020. The latter nonrejection is consistent with our previous observation that the volatility of SPY was high even before the announcement. Meanwhile, the test also strongly rejects the null hypothesis of distributional continuity for the volume series, echoing the burst of trading activity seen in Figure 2.

The permutation test applied to the two illiquidity measures generates mixed results. We find some moderate evidence for the distributional discontinuity of Amihud's measure shortly after the March 3 and December 16 announcements in 2020. For the case with a 30-minute window, we do not reject the null of continuity for any of the announcements. The overall evidence suggests that the FOMC announcements under study did not lead to an abrupt change in the market impact coefficient gauged by Amihud's measure. Needless to say, this finding per se does not imply that the liquidity condition is unchanged after the announcement, as the notion of liquidity is a multifaceted concept. Indeed, we see that the permutation test applied to the bid-ask spread always strongly rejects the null of continuity. The post-announcement spread tends to be larger than its pre-announcement level, suggesting that it is significantly more costly to trade during the post-announcement trading session.

All in all, the empirical illustration above demonstrates how the proposed permutation test may be used to test for distributional discontinuities for a variety of market variables. This type of versatility is not easily attained by existing methods in the high-frequency econometrics literature. We also see that interesting empirical findings may

[16]Prior work on volatility jumps (see, e.g., Bollerslev, Li, and Xue (2018) and Li, Todorov, and Zhang (2021)) focuses exclusively on the case in which volatility is the scaling factor of Brownian shocks. Our permutation test is related to the prior work, but is valid under a more general notion of volatility with respect to non-Brownian shocks.

be obtained even with a small number of observations, which confirms the practical relevance of allowing the k_n window to be possibly fixed in our asymptotic theory for the permutation test.

5. Concluding remarks

In this paper, we propose using a permutation test to detect discontinuities in an economic model at a cutoff point. Relative to the existing literature, we show that the permutation test is well suited for event studies based on time-series data. While nonparametric t-tests have been widely used for this purpose in various empirical contexts, the permutation test proposed in this paper provides a distinct alternative. Instead of relying on asymptotic (mixed) Gaussianity from central limit theorems, we exploit finite-sample properties of the permutation test in the approximating, or "coupling," two-sample problem.

We demonstrate that our new theory is broadly useful in a wide range of problems in the infill asymptotic time-series setting, which justifies using the permutation test to detect jumps in economic variables such as volatility, trading activity, and liquidity. Compared with the conventional nonparametric t-test, the proposed permutation test has several distinct features. First, the permutation test provides asymptotic size control regardless of whether the sizes of the local subsamples are fixed or growing to infinity. In the latter case, we also establish that the permutation test is consistent. Second, the permutation test is versatile, as it can be applied without modification to many different contexts and under relatively weak conditions.

Appendix A: Proofs

Throughout the proofs, we use K to denote a positive constant that may change from line to line, and write K_p to emphasize its dependence on some parameter p. For any event $E \in \mathcal{F}$, we identify it with the associated indicator random variable.

Proof of Theorem 2.1. Step 1. Define ϕ_n in the same way as $\hat{\phi}_n$ but with $(Y_{n,i})_{i\in\mathcal{I}_n}$ replaced by $(U_{n,i})_{i\in\mathcal{I}_n}$. In this step, we show that

$$\mathbb{E}[\hat{\phi}_n] = \mathbb{E}[\phi_n] + o(1). \tag{11}$$

Let $\tilde{\phi}_n$ be defined in the same way as $\hat{\phi}_n$, but with $(Y_{n,i})_{i\in\mathcal{I}_n}$ replaced by $(\widetilde{Y}_{n,i})_{i\in\mathcal{I}_n}$, as defined in Assumption 2.1(iii). Since $(\widetilde{Y}_{n,i})_{i\in\mathcal{I}_n}$ and $(Y_{n,i})_{i\in\mathcal{I}_n}$ have the same (conditional) distribution,

$$\mathbb{E}[\tilde{\phi}_n] = \mathbb{E}[\hat{\phi}_n]. \tag{12}$$

Let $E_n \in \mathcal{F}$ be the event where the ordered values of $(U_{n,i})_{i\in\mathcal{I}_n}$ and $(\widetilde{Y}_{n,i})_{i\in\mathcal{I}_n}$ correspond to the same permutation of $\mathcal{I}_n$. Since the test statistic is only a function of the rank of the observations, we have $\tilde{\phi}_n = \phi_n$ in restriction to E_n. Hence,

$$\left|\mathbb{E}[\tilde{\phi}_n] - \mathbb{E}[\phi_n]\right| = \left|\mathbb{E}\left[\tilde{\phi}_n E_n^c\right] - \mathbb{E}\left[\phi_n E_n^c\right]\right| \le \mathbb{P}\left(E_n^c\right). \tag{13}$$

By (12) and (13), (11) follows from $\mathbb{P}(E_n^c) = o(1)$, which will be proved below.

Let $A_{n,i,j} \equiv \{U_{n,j} - U_{n,i} \geq 0, \widetilde{Y}_{n,j} - \widetilde{Y}_{n,i} < 0\}$ for every $(i,j) \in \mathcal{I}_n \times \mathcal{I}_n$, and note that $E_n^c \subseteq \cup_{i,j} A_{n,i,j}$. Recall the elementary fact that if a sequence of random variables $X_n = o_p(1)$, then there exists a real sequence $\delta_n = o(1)$ such that $\mathbb{P}(|X_n| \leq \delta_n) \to 1$. Under Assumption 2.1(iii), by applying this result to $X_n = 2\max_{i\in\mathcal{I}_n} |\widetilde{Y}_{n,i} - U_{n,i}|k_n^2$, we can find a sequence $\delta_n = o(1)$ such that

$$\mathbb{P}\Big(\max_{i\in\mathcal{I}_n} |\widetilde{Y}_{n,i} - U_{n,i}| \leq \delta_n k_n^{-2}/2\Big) \to 1. \tag{14}$$

We then observe that

$$\begin{aligned} A_{n,i,j} &\subseteq \big\{U_{n,j} - U_{n,i} \geq \delta_n k_n^{-2}, \widetilde{Y}_{n,j} - \widetilde{Y}_{n,i} < 0\big\} \cup \big\{0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}\big\} \\ &\subseteq \big\{\big|\widetilde{Y}_{n,j} - \widetilde{Y}_{n,i} - (U_{n,j} - U_{n,i})\big| > \delta_n k_n^{-2}\big\} \cup \big\{0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}\big\} \\ &\subseteq \Big\{\max_{i\in\mathcal{I}_n} |\widetilde{Y}_{n,i} - U_{n,i}| > \delta_n k_n^{-2}/2\Big\} \cup \big\{0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}\big\}. \end{aligned}$$

Therefore,

$$E_n^c \subseteq \cup_{i,j} A_{n,i,j} \subseteq \Big\{\max_{i\in\mathcal{I}_n} |\widetilde{Y}_{n,i} - U_{n,i}| > \delta_n k_n^{-2}/2\Big\} \cup \big(\cup_{i,j\in\mathcal{I}_n} \big\{0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}\big\}\big),$$

which, together with (14), implies that

$$\mathbb{P}(E_n^c) \leq \mathbb{P}\big(\cup_{i,j\in\mathcal{I}_n} \big\{0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}\big\}\big) + o(1). \tag{15}$$

Next, consider the following argument:

$$\begin{aligned} \mathbb{P}\big(\cup_{i,j\in\mathcal{I}_n} \big\{0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}\big\}|\mathcal{G}_n\big) &\leq \sum_{i,j\in\mathcal{I}_n} \mathbb{P}\big(0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}|\mathcal{G}_n\big) \\ &\leq 2k_n \sum_{i\in\mathcal{I}_n} \sup_{x\in\mathbb{R}} \mathbb{P}\big(|U_{n,i} - x| \leq \delta_n k_n^{-2}|\mathcal{G}_n\big) \\ &= O_p(\delta_n) = o_p(1), \end{aligned} \tag{16}$$

where the last line holds by Assumption 2.1(ii). By (16) and the bounded convergence theorem,

$$\mathbb{P}\big(\cup_{i,j\in\mathcal{I}_n} \big\{0 \leq U_{n,j} - U_{n,i} < \delta_n k_n^{-2}\big\}\big) = o(1). \tag{17}$$

By combining (15) and (17), we conclude that $\mathbb{P}(E_n^c) = o(1)$, as desired.

Step 2. We now prove the assertions in parts (a) and (b) of the theorem. In view of (11), we only need to prove $\mathbb{E}[\phi_n] \to \alpha$ and $\mathbb{E}[\phi_n] \to 1$ in these two parts, respectively. For part (a), note that $(U_{n,i})_{i\in\mathcal{I}_n}$ are conditionally i.i.d. and so permutations constitute a group of transformations that satisfy the randomization hypothesis in Lehmann and Romano (2005, Definition 15.2.1). Then Lehmann and Romano (2005, Theorem 15.2.1) imply that $\mathbb{E}[\phi_n|\,\mathcal{G}_n] = \alpha$, and $\mathbb{E}[\phi_n] = \alpha$ then follows from the law of iterated expectations.

To prove part (b), we need some additional notation. To emphasize the dependence of $\widehat{T}_n$, $\widehat{T}_n^*$, and $\hat{\phi}_n$ on the original data $(Y_{n,i})_{i\in\mathcal{I}_n}$, we explicitly write them as $\widehat{T}_n(Y)$, $\widehat{T}_n^*(Y)$,

and $\hat{\phi}_n(Y)$. With this notation, we can rewrite $\phi_n = \hat{\phi}_n(U)$, since it is computed in the same way as $\hat{\phi}_n$ but with $(Y_{n,i})_{i\in\mathcal{I}_n}$ replaced by $(U_{n,i})_{i\in\mathcal{I}_n}$.

We first analyze the asymptotic behavior of $\widehat{T}_n(U)$. Define the empirical analogue of $Q_{j,n}(\cdot)$ as

$$\widehat{Q}_{j,n}(x) \equiv \frac{1}{k_n}\sum_{i\in\mathcal{I}_{j,n}} 1\{U_{n,i} \le x\}.$$

Since the variables $(U_{n,i})_{i\in\mathcal{I}_{j,n}}$ are $\mathcal{G}_n$-conditionally i.i.d.,

$$\mathbb{E}\big[\big(\widehat{Q}_{j,n}(x) - Q_{j,n}(x)\big)^2 | \mathcal{G}_n\big] \le O\big(k_n^{-1}\big) = o(1).$$

By Markov's inequality and the law of iterated expectations, this implies that $\widehat{Q}_{j,n}(x) - Q_{j,n}(x) = o_p(1)$ for each $x \in \mathbb{R}$. This and a classical Glivenko–Cantelli theorem (e.g., Davidson (1994, Theorem 21.5)) imply that

$$\sup_{x\in\mathbb{R}} \big|\widehat{Q}_{j,n}(x) - Q_{j,n}(x)\big| = o_p(1), \quad \text{for } j \in \{1,2\}. \tag{18}$$

By definition,

$$\widehat{T}_n(U) \equiv \frac{1}{2k_n}\sum_{i\in\mathcal{I}_n}\big(\widehat{Q}_{1,n}(U_{n,i}) - \widehat{Q}_{2,n}(U_{n,i})\big)^2.$$

In addition, we define

$$S_n \equiv \frac{1}{2k_n}\sum_{i\in\mathcal{I}_n}\big(Q_{1,n}(U_{n,i}) - Q_{2,n}(U_{n,i})\big)^2.$$

Note that the functions $\widehat{Q}_{j,n}(\cdot)$ and $Q_{j,n}(\cdot)$ are uniformly bounded. Hence, by the triangle inequality and (18),

$$\begin{aligned}\big|\widehat{T}_n(U) - S_n\big| &\le \frac{1}{2k_n}\sum_{i\in\mathcal{I}_n}\big|\big(\widehat{Q}_{1,n}(U_{n,i}) - \widehat{Q}_{2,n}(U_{n,i})\big)^2 - \big(Q_{1,n}(U_{n,i}) - Q_{2,n}(U_{n,i})\big)^2\big| \\ &\le \frac{K}{k_n}\sum_{i\in\mathcal{I}_n}\sum_{j\in\{1,2\}}\big|\widehat{Q}_{j,n}(U_{n,i}) - Q_{j,n}(U_{n,i})\big| = o_p(1). \end{aligned} \tag{19}$$

Conditional on $\mathcal{G}_n$, the bounded random functions $Q_{1,n}(\cdot)$ and $Q_{2,n}(\cdot)$ can be treated as deterministic functions. Next, note that

$$\begin{aligned} S_n &= \frac{1}{2}\sum_{j\in\{1,2\}}\int\big(Q_{1,n}(x) - Q_{2,n}(x)\big)^2 dQ_{j,n}(x) + o_p(1) \\ &= \int\big(Q_{1,n}(x) - Q_{2,n}(x)\big)^2 d\overline{Q}_n(x) + o_p(1), \end{aligned} \tag{20}$$

where the first equality holds by a law of large numbers for the conditionally i.i.d. variables $(U_{n,i})_{i\in\mathcal{I}_{j,n}}$ for $j = 1, 2$, and the second equality holds by the definition of $\overline{Q}_n$. By

combining (19) and (20), we deduce that

$$\widehat{T}_n(U) = \int \big(Q_{1,n}(x) - Q_{2,n}(x)\big)^2 \, d\overline{Q}_n(x) + o_p(1). \tag{21}$$

In turn, by (21) and the condition in part (b), we conclude that

$$\mathbb{P}\big(\widehat{T}_n(U) > \delta_n\big) \to 1 \quad \text{for any } \delta_n = o(1). \tag{22}$$

Next, we analyze the asymptotic behavior of $\widehat{T}_n^*(U)$. It is useful to consider the following representation of this variable. We denote $U_{\tilde{\pi}} = (U_{n,\tilde{\pi}(i)})_{i\in\mathcal{I}_n}$, where $\tilde{\pi}$ is a random permutation of $\mathcal{I}_n$, independent from the data, and is drawn uniformly from the set of all permutations of $\mathcal{I}_n$. By definition, $\widehat{T}_n^*(U)$ is the $1-\alpha$ quantile of $\widehat{T}_n(U_{\tilde{\pi}})$, conditional on the sample, where the randomness comes from the random realization of $\tilde{\pi}$. To analyze the permutation distribution, we construct an additional coupling sequence of $(U_{n,i})_{i\in\mathcal{I}_n}$ following the method of Chung and Romano (2013, Section 5.3). We note that their coupling construction does not require the null hypothesis to hold, and it is thus suitable for our current purposes. The result of their coupling construction is another random sequence $(U'_{n,i})_{i\in\mathcal{I}_n}$ such that (i) $U_{n,i} = U'_{n,i}$ for all i in some random subset $\mathcal{I}'_n \subseteq \mathcal{I}_n$; (ii) the cardinality of $\mathcal{I}_n \setminus \mathcal{I}'_n$, denoted D_n, satisfies $\mathbb{E}[D_n] = O(k_n^{1/2})$; and (iii) $(U'_{n,i})_{i\in\mathcal{I}_n}$ are $\mathcal{G}_n$-conditionally i.i.d. with marginal distribution $\overline{Q}_n$.

For any fixed arbitrary permutation π and for $j \in \{1,2\}$, define

$$\widehat{Q}_{j,n}(x;\pi) \equiv \frac{1}{k_n}\sum_{i\in\mathcal{I}_{j,n}} 1\{U_{n,\pi(i)} \le x\} \quad \text{and} \quad \widehat{Q}'_{j,n}(x;\pi) \equiv \frac{1}{k_n}\sum_{i\in\mathcal{I}_{j,n}} 1\big\{U'_{n,\pi(i)} \le x\big\}.$$

By repeatedly using the triangle inequality,

$$\begin{aligned}
\big|\widehat{T}_n(U_\pi) - \widehat{T}_n\big(U'_\pi\big)\big| &= \frac{1}{2k_n}\Bigg|\sum_{i\in\mathcal{I}_n}\Big(\big(\widehat{Q}_{1,n}(U_{n,\pi(i)};\pi) - \widehat{Q}_{2,n}(U_{n,\pi(i)};\pi)\big)^2 \\
&\qquad - \big(\widehat{Q}'_{1,n}\big(U'_{n,\pi(i)};\pi\big) - \widehat{Q}'_{2,n}\big(U'_{n,\pi(i)};\pi\big)\big)^2\Big)\Bigg| \\
&\le \frac{K}{k_n}\sum_{j\in\{1,2\}}\sum_{i\in\mathcal{I}_n}\big|\widehat{Q}_{j,n}(U_{n,\pi(i)};\pi) - \widehat{Q}'_{j,n}\big(U'_{n,\pi(i)};\pi\big)\big| \\
&\le \frac{K}{k_n^2}\sum_{i,k\in\mathcal{I}_n}\big|1\{U_{n,\pi(k)} \le U_{n,\pi(i)}\} - 1\big\{U'_{n,\pi(k)} \le U'_{n,\pi(i)}\big\}\big| \\
&\le KD_n/k_n = o_p(1),
\end{aligned} \tag{23}$$

where the last inequality uses the fact that $(U_{n,i}, U_{n,k}) = (U'_{n,i}, U'_{n,k})$ if $(i,k) \in \mathcal{I}'_n \times \mathcal{I}'_n$, and so the summation on the previous line only has $(2k_n)^2 - (2k_n - D_n)^2 \le 4k_nD_n$ bounded terms that can be different from zero, and the $o_p(1)$ statement follows from $\mathbb{E}[D_n] = O(k_n^{1/2})$, $k_n \to \infty$, and Markov's inequality.

For any fixed arbitrary permutation π, $\widehat{T}_n(U'_\pi)$ is the Cramér–von Mises statistic for the $\mathcal{G}_n$-conditionally i.i.d. variables $(U'_{n,\pi(i)})_{i\in\mathcal{I}_n}$. Hence, by a similar argument leading

to (22), we have $\widehat{T}_n(U'_\pi) = o_p(1)$. By combining this with (23), it follows that

$$\widehat{T}_n(U_\pi) = o_p(1). \tag{24}$$

Since this result holds for any arbitrary fixed permutation π, it also holds for any pair of permutations considered at random from the set of all possible permutations of $\mathcal{I}_n$, independently from the data. By elementary properties of stochastic convergence, this implies the so-called Hoeffding's condition (e.g., Lehmann and Romano (2005, equation (15.10))). By this and Lehmann and Romano (2005, Theorem 15.2.3), the permutation distribution associated with the test statistic $\widehat{T}_n(U)$, conditional on the data, converges to zero in probability. As a corollary of this,

$$\widehat{T}^*_n(U) = o_p(1). \tag{25}$$

From (22) and (25), it is easy to see that $\widehat{T}_n(U) > \widehat{T}^*_n(U)$ with probability approaching 1. This further implies that $\mathbb{E}[\phi_n] \to 1$, which together with (11) proves the assertion of part (b). □

Proof of Theorem 2.2. (a) We prove the assertion of part (a) by applying Theorem 2.1(a). We construct the coupling variable $U_{n,i}$ as follows:

$$U_{n,i} = g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i}), \quad \text{for all } i \in \mathcal{I}_n. \tag{26}$$

We set $\mathcal{G}_n = \mathcal{F}_{(i^*-k_n)\Delta_n}$. By Assumption 2.2, $(\epsilon_{n,i})_{i\in\mathcal{I}_n}$ are i.i.d. and independent of $\mathcal{G}_n$. Since $\zeta_{(i^*-k_n)\Delta_n}$ is $\mathcal{G}_n$-measurable, the variables $(U_{n,i})_{i\in\mathcal{I}_n}$ are $\mathcal{G}_n$-conditionally i.i.d. This verifies the condition in part (a) of Theorem 2.1, which also implies Assumption 2.1(i). It remains to verify conditions (ii) and (iii) in Assumption 2.1.

By a standard localization argument (see Jacod and Protter (2012, Section 4.4.1)), we can strengthen Assumption 2.3 by assuming $T_1 = \infty$, $\mathcal{K}_m = \mathcal{K}$, and $K_m = K$ for some fixed compact set $\mathcal{K}$ and constant $K > 0$. In particular, $\zeta_{(i^*-k_n)\Delta_n}$ takes values in the compact set $\mathcal{K}$. By Assumption 2.2, it is then easy to see that the $\mathcal{G}_n$-conditional probability density of $U_{n,i} = g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i})$ is uniformly bounded (and it does not depend on i). This implies condition (ii) of Assumption 2.1.

Finally, we verify condition (iii) of Assumption 2.1. By Assumption 2.2(i), for each $i \in \mathcal{I}_n$, $\varepsilon_{n,i}$ is independent of $\mathcal{F}_{i\Delta_n}$. Since $\zeta_{i\Delta_n}$ and $\zeta_{(i^*-k_n)\Delta_n}$ are $\mathcal{F}_{i\Delta_n}$-measurable, we deduce from Assumption 2.3(i) that

$$\mathbb{E}\big[\big|g(\zeta_{i\Delta_n}, \epsilon_{n,i}) - g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i})\big|^2 \big| \mathcal{F}_{i\Delta_n}\big] \le K a_n^2 \|\zeta_{i\Delta_n} - \zeta_{(i^*-k_n)\Delta_n}\|^2. \tag{27}$$

Note that under the null hypothesis with $\Delta\zeta_\tau = 0$, the processes ζ_t and $\tilde{\zeta}_t$ are identical. Hence, by Assumption 2.3(ii) and (27),

$$\big\|g(\zeta_{i\Delta_n}, \epsilon_{n,i}) - g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i})\big\|_2 \le K a_n k_n^{1/2} \Delta_n^{1/2}.$$

By the maximal inequality under the L_2 norm (see, e.g., Vaart and Wellner (1996, Lemma 2.2.2)), we further deduce that

$$\Big\|\max_{i\in\mathcal{I}_n} \big|g(\zeta_{i\Delta_n}, \epsilon_{n,i}) - g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i})\big|\Big\|_2 \le K a_n k_n \Delta_n^{1/2}. \tag{28}$$

Recall that $a_n k_n^3 \Delta_n^{1/2} = o(1)$ by assumption. Hence,

$$\max_{i \in \mathcal{I}_n} \left| g(\zeta_{i\Delta_n}, \epsilon_{n,i}) - g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i}) \right| = o_p\big(k_n^{-2}\big). \tag{29}$$

Note that, by the definitions in (4) and (26),

$$Y_{n,i} - U_{n,i} = g(\zeta_{i\Delta_n}, \epsilon_{n,i}) - g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i}) + R_{n,i}. \tag{30}$$

Combining (29), (30), and Assumption 2.3(iii), we deduce that $\max_{i \in \mathcal{I}_n} |Y_{n,i} - U_{n,i}| = o_p(k_n^{-2})$, which verifies Assumption 2.1(iii). We have now verified all the conditions needed in Theorem 2.1(a), which proves the assertion of part (a) of Theorem 2.2.

(b) We prove the assertion of part (b) by applying Theorem 2.1(b). Under the maintained alternative hypothesis, we have $\Delta\zeta_\tau = c$ for some constant $c \neq 0$. The coupling variable now takes the following form:

$$U_{n,i} = \begin{cases} g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i}) & i \in \mathcal{I}_{1,n}, \\ g(\zeta_{(i^*-k_n)\Delta_n} + c, \epsilon_{n,i}) & i \in \mathcal{I}_{2,n}. \end{cases} \tag{31}$$

Under Assumption 2.2, it is easy to see that, for each $j \in \{1, 2\}$, the variables $(U_{n,i})_{i \in \mathcal{I}_{j,n}}$ are $\mathcal{G}_n$-conditionally i.i.d., which verifies Assumption 2.1(i).

We now turn to the remaining conditions in Assumption 2.1. As in part (a), we can invoke the standard localization procedure and assume that the ζ_t process takes value in a compact set $\mathcal{K}$. Note that

$$\zeta_\tau - (\zeta_{(i^*-k_n)\Delta_n} + \Delta\zeta_\tau) = \zeta_{\tau-} - \zeta_{(i^*-k_n)\Delta_n} = o_p(1),$$

where the $o_p(1)$ statement follows from the fact that the ζ_t process is càdlàg and $k_n\Delta_n \to 0$. Therefore, by enlarging the compact set $\mathcal{K}$ slightly if necessary, we also have $\zeta_{(i^*-k_n)\Delta_n} + c \in \mathcal{K}$ with probability approaching 1. Then we can verify Assumption 2.1(ii) following the same argument as in part (a). The verification of Assumption 2.1(iii) is also similar.

Finally, we verify the condition in Theorem 2.1(b) pertaining to the conditional CDFs. Note that

$$Q_{1,n}(x) = F_{\zeta_{(i^*-k_n)\Delta_n}}(x) \quad \text{and} \quad Q_{2,n}(x) = F_{\zeta_{(i^*-k_n)\Delta_n}+c}(x).$$

It is then easy to see that

$$2\int \big(Q_{1,n}(x) - Q_{2,n}(x)\big)^2 \, d\overline{Q}_n(x) \geq \int \big(F_{\zeta_{(i^*-k_n)\Delta_n}}(x) - F_{\zeta_{(i^*-k_n)\Delta_n}+c}(x)\big)^2 \, dF_{\zeta_{(i^*-k_n)\Delta_n}}(x).$$

Since $\zeta_{(i^*-k_n)\Delta_n}$ takes values in the compact set $\mathcal{K}$, Assumption 2.2(iii) implies that the lower bound in the above display is bounded away from zero. Hence, $\int (Q_{1,n}(x) - Q_{2,n}(x))^2 \, d\overline{Q}_n(x) > \delta_n$ for any real sequence $\delta_n = o(1)$. We have now verified all conditions for Theorem 2.1(b), which proves the assertion of part (b) of Theorem 2.2. □

PROOF OF THEOREM 2.3. This proof follows from similar arguments to those used to prove Theorem 2.1. For the sake of brevity, we focus on the only substantial difference, which is how we establish that $\mathbb{P}(E_n^c) = o(1)$. Recall that E_n denotes the event where the ordered values of $(U_{n,i})_{i\in\mathcal{I}_n}$ and $(\widetilde{Y}_{n,i})_{i\in\mathcal{I}_n}$ correspond to the same permutation of $\mathcal{I}_n$. In the case of this proof, this result follows from

$$\mathbb{P}\big(E_n^c\big) \le \mathbb{P}\big(\cup_{i\in\mathcal{I}_n}\{\widetilde{Y}_{n,i} \neq U_{n,i}\}\big) \le \sum_{i\in\mathcal{I}_n} \mathbb{P}(\widetilde{Y}_{n,i} \neq U_{n,i}) = o(1),$$

where the first inequality follows from $E_n^c \subseteq \cup_{i\in\mathcal{I}_n}\{\widetilde{Y}_{n,i} \neq U_{n,i}\}$ and the convergence follows from the assumption that $\mathbb{P}(\widetilde{Y}_{n,i} \neq U_{n,i}) = o(k_n^{-1})$ uniformly in $i \in \mathcal{I}_n$. □

PROOF OF THEOREM 2.4. (a) We prove this assertion by applying Theorem 2.3(a). We shall verify the conditions in Theorem 2.3 for $\widetilde{Y}_{n,i} = Y_{n,i}$, $U_{n,i} = g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i})$, and $\mathcal{G}_n = \mathcal{F}_{(i^*-k_n)\Delta_n}$. By assumption, the variables $(\epsilon_{n,i})_{i\in\mathcal{I}_n}$ are i.i.d. and independent of $\mathcal{G}_n$. Hence, the variables $(U_{n,i})_{i\in\mathcal{I}_n}$ are $\mathcal{G}_n$-conditionally i.i.d.

It remains to verify that $\mathbb{P}(Y_{n,i} \neq U_{n,i}) = o(k_n^{-1})$ uniformly in $i \in \mathcal{I}_n$. By repeating the localization argument used in the proof of Theorem 2.2, we can strengthen Assumption 2.4 with $T_1 = \infty$ without loss of generality. In particular, ζ_t takes values in some compact subset $\mathcal{K} \subseteq \mathcal{Z}$. Note that for each $i \in \mathcal{I}_n$, $\epsilon_{n,i}$ is independent of $(\zeta_{i\Delta_n}, \zeta_{(i^*-k_n)\Delta_n})$. By Assumption 2.4(i), we thus have $\mathbb{P}(Y_{n,i} \neq U_{n,i}|\mathcal{G}_n) \le K\|\zeta_{i\Delta_n} - \zeta_{(i^*-k_n)\Delta_n}\|$. Then, by Assumption 2.4(ii), we further have $\mathbb{P}(Y_{n,i} \neq U_{n,i}) \le K(k_n\Delta_n)^{1/2}$. The condition $\mathbb{P}(Y_{n,i} \neq U_{n,i}) = o(k_n^{-1})$ then follows from $k_n^3\Delta_n = o(1)$. By Theorem 2.3(a), we have $\mathbb{E}[\hat{\phi}_n] \to \alpha$ as asserted.

(b) We prove this assertion by applying Theorem 2.3(b). We verify the conditions in Theorem 2.3 for $\widetilde{Y}_{n,i} = Y_{n,i}$, $\mathcal{G}_n = \mathcal{F}_{(i^*-k_n)\Delta_n}$, and

$$U_{n,i} = \begin{cases} g(\zeta_{(i^*-k_n)\Delta_n}, \epsilon_{n,i}) & \text{if } i \in \mathcal{I}_{1,n}, \\ g(\zeta_{(i^*-k_n)\Delta_n} + c, \epsilon_{n,i}) & \text{if } i \in \mathcal{I}_{2,n}. \end{cases}$$

Following the same argument as in part (a), we see that $(U_{n,i})_{i\in\mathcal{I}_{j,n}}$ are $\mathcal{G}_n$-conditionally i.i.d. for each $j \in \{1,2\}$, and $\mathbb{P}(Y_{n,i} \neq U_{n,i}) = o(k_n^{-1})$ uniformly in $i \in \mathcal{I}_n$. Assumption 2.2(iii) also ensures that $\mathbb{P}(\int (Q_{1,n}(x) - Q_{2,n}(x))^2 \, d\overline{Q}_n(x) > \delta_n) \to 1$ for any real sequence $\delta_n = o(1)$. By Theorem 2.3(b), we have that $\mathbb{E}[\hat{\phi}_n] \to 1$, as asserted. □

APPENDIX B: EXTENSION TO OTHER TEST STATISTICS

As explained in Remark 2.2, our main results extend beyond the Cramér–von Mises statistic in (1). We characterize the relevant class of test statistics by the following high-level assumption.

ASSUMPTION B.1. *$\widehat{T}_n \equiv \Psi_n((Y_{n,i})_{i\in\mathcal{I}_n})$, where $(\Psi_n)_{n\in\mathbb{N}}$ is a sequence of functions that satisfies the following conditions*:

(a) $\widehat{T}_n$ is a rank statistic, that is, for any $(Y_{n,i})_{i\in\mathcal{I}_n}$ and $(Y'_{n,i})_{i\in\mathcal{I}_n}$ with $\operatorname{sign}(Y_{n,i} - Y_{n,j}) = \operatorname{sign}(Y'_{n,i} - Y'_{n,j})$ for all $i, j \in \mathcal{I}_n$, $\Psi_n((Y_{n,i})_{i\in\mathcal{I}_n}) = \Psi_n((Y'_{n,i})_{i\in\mathcal{I}_n})$.

(b) For any $(Y_{n,i})_{i\in\mathcal{I}_n}$ and $(Y'_{n,i})_{i\in\mathcal{I}_n}$, $|\Psi_n((Y_{n,i})_{i\in\mathcal{I}_n}) - \Psi_n((Y'_{n,i})_{i\in\mathcal{I}_n})| = O_p(D_n/k_n)$ with $D_n = |\{i \in \mathcal{I}_n : Y_{n,i} \neq Y'_{n,i}\}|$.

Assumption B.1 is satisfied for a large class of test statistics, which includes the Cramér–von Mises and Kolmogorov–Smirnov statistics. In fact, the proof of Theorem 2.1 shows that the Cramér–von Mises statistic satisfies Assumption B.1, and an analogous argument can be used to extend this to the Kolmogorov–Smirnov statistic. We now briefly describe the assumption. Assumption B.1(a) is an essential ingredient to our methodology. In turn, Assumption B.1(b) is a mild regularity condition that limits the influence that a few sample observations can have on the test statistic, and is only required to establish our consistency result.

The result proves Theorem 2.1 for any test statistic that satisfies Assumption B.1. Since Theorem 2.1 is the key to all of the results in the paper, this effectively implies that our findings extend to the class of statistics characterized by Assumption B.1, as claimed in Remark 2.2.

THEOREM B.1. *Under Assumptions 2.1 and B.1 (instead of (1)),*

(a) If the variables $(U_{n,i})_{i\in\mathcal{I}_n}$ have the same $\mathcal{G}_n$-conditional distribution, we have $\mathbb{E}[\hat{\phi}_n] \to \alpha$.

(b) Let $\widehat{T}_n(U)$ denote the test statistic but applied to $(U_{n,i})_{i\in\mathcal{I}_n}$ instead of $(Y_{n,i})_{i\in\mathcal{I}_n}$. If $k_n \to \infty$ and $\mathbb{P}(\widehat{T}_n(U) > \delta_n) \to 1$ for any real sequence $\delta_n = o(1)$, we have $\mathbb{E}[\hat{\phi}_n] \to 1$.

PROOF. This proof follows closely that of Theorem 2.1, which has two steps. Step 1 remains unchanged, as it only relies on Assumption 2.1 and the fact that $\widehat{T}_n$ is a rank statistic, imposed in Assumption B.1(a). Part (a) of Step 2 also remains unchanged, as it is entirely based on Step 1. To complete this proof, it then suffices to cover the analog of part (b) of Step 2.

We begin by considering the asymptotic behavior of $\widehat{T}_n^*(U)$, that is, the $1-\alpha$ quantile of $\widehat{T}_n(U_{\tilde{\pi}})$, conditional on the sample, where the randomness comes from the realization of $\tilde{\pi}$. As in the proof of Theorem 2.1, we rely on the coupling construction based on Chung and Romano (2013, Section 5.3), which produces a random sequence $(U'_{n,i})_{i\in\mathcal{I}_n}$ such that (i) $U_{n,i} = U'_{n,i}$ for all i in some random subset $\mathcal{I}'_n \subseteq \mathcal{I}_n$; (ii) $D_n = |\{i \in \mathcal{I}_n : U_{n,i} \neq U'_{n,i}\}|$ satisfies $\mathbb{E}[D_n] = O(k_n^{1/2})$; and (iii) $(U'_{n,i})_{i\in\mathcal{I}_n}$ are $\mathcal{G}_n$-conditionally i.i.d. Then, for any fixed arbitrary permutation π,

$$\begin{aligned}\left|\widehat{T}_n(U_\pi) - \widehat{T}_n\big(U'_\pi\big)\right| &= \left|\Psi_n\big((U_{n,\pi(i)})_{i\in\mathcal{I}_n}\big) - \Psi_n\big(\big(U'_{n,\pi(i)}\big)_{i\in\mathcal{I}_n}\big)\right| \\ &= O_p\big(D_n/k_n^{1/2+\varepsilon}\big) = o_p(1),\end{aligned} \tag{32}$$

where the second equality relies on $|\{i \in \mathcal{I}_n : U_{n,\pi(i)} \neq U'_{n,\pi(i)}\}| \leq D_n$ and Assumption B.1(b), and the last equality relies on $\mathbb{E}[D_n] = O(k_n^{1/2})$, $k_n \to \infty$, and Markov's inequality. We can then repeat the arguments in the proof of Theorem 2.1 to conclude that $\widehat{T}_n(U'_\pi) = o_p(1)$ for any fixed arbitrary permutation π. By this with (32), we conclude that $\widehat{T}_n(U_\pi) = o_p(1)$. Since π was arbitrarily chosen, it then follows that

$$\widehat{T}_n^*(U) = o_p(1). \tag{33}$$

From the assumption in part (b) and (33), it is easy to see that $\widehat{T}_n(U) > \widehat{T}_n^*(U)$ with probability approaching one. We can then repeat the remaining arguments in the proof of Theorem 2.1 to complete this proof. □

References

Amihud, Yakov (2002), "Illiquidity and stock returns: Cross-section and time-series effects." *Journal of Financial Markets*, 5, 31–56. [57]

Andersen, Torben G. (1996), "Return volatility and trading volume: An information flow interpretation of stochastic volatility." *Journal of Finance*, 51, 169–204. [52]

Andrews, Donald W. K. (1993), "Tests for parameter instability and structural change with unknown change point." *Econometrica*, 61, 821–856. [38]

Antoch, Jaromír and Maria Hušková (2001), "Permutation tests in change point analysis." *Statistics & Probability Letters*, 53, 273–291. [39]

Bai, Jushan and Pierre Perron (1998), "Estimating and testing linear models with multiple structural changes." *Econometrica*, 66, 47–78. [38]

Barndorff-Nielsen, Ole E. and Neil Shephard (2001), "Non-Gaussian Ornstein–Uhlenbeck-based models and some of their uses in financial economics." *Journal of the Royal Statistical Society, Series B*, 63, 167–241. [40]

Barndorff-Nielsen, Ole E. and Neil Shephard (2006), "Econometrics of testing for jumps in financial economics using bipower variation." *Journal of Financial Econometrics*, 4, 1–30. [38]

Bollerslev, Tim (2011), "Estimation of jump tails." *Review of Economic Studies*, 85, 2005–2041. [54]

Bollerslev, Tim, Jia Li, and Yuan Xue (2018), "Volume, volatility, and public news announcements." *Review of Economic Studies*, 85, 2005–2041. [38, 40, 46, 54, 59]

Bugni, Federico A. and Ivan A. Canay (2021), "Testing continuity of a density via g-order statistics in the regression discontinuity design." *Journal of Econometrics*, 221, 138–159. [40]

Bugni, Federico A., Jia Li, and Qiyuan Li (2023), "Supplement to 'Permutation-based tests for discontinuities in event studies'." *Quantitative Economics Supplemental Material*, 14, https://doi.org/10.3982/QE1775. [37]

Calonico, Sebastian, Matias D. Cattaneo, and Rocio Titiunik (2014), "Robust nonparametric confidence intervals for regression-discontinuity designs." *Econometrica*, 82, 2295–2326. [38]

Canay, Ivan A. and Vishal Kamat (2017), "Approximate permutation tests and induced order statistics in the regression discontinuity design." *The Review of Economic Studies*, 85, 1577–1608. [38, 40, 44, 45]

Cattaneo, Matias D., Brigham R. Frandsen, and Rocio Titiunik (2015), "Randomization inference in the regression discontinuity design: An application to party advantages in the US senate." *Journal of Causal Inference*, 3, 1–24. [39, 40]

Cattaneo, Matias D., Rocio Titiunik, and Gonzalo Vazquez-Bare (2017), "Comparing inference approaches for RD designs: A reexamination of the effect of head start on child mortality." *Journal of Policy Analysis and Management*, 36, 643–681. [38, 40]

Chernozhukov, Victor, Kaspar Wuthrich, and Yinchu Zhu (2018), *Exact and Robust Conformal Inference Methods for Predictive Machine Learning With Dependent Data*. Technical report. arXiv:1802.06300v3. [39]

Chow, Gregory C. (1960), "Tests of equality between sets of coefficients in two linear regressions." *Econometrica*, 28, 591–605. [37]

Chung, EunYi and Joseph P. Romano (2013), "Exact and asymptotically robust permutation tests." *Annals of Statistics*, 41, 484–507. [39, 45, 63, 67]

Cochrane, John H. and Monika Piazzesi (2002), "The fed and interest rates—a high-frequency identification." *American Economic Review*, 92, 90–95. [39, 40]

Comte, Fabienne and Eric Renault (1998), "Long memory in continuous time stochastic volatility models." *Mathematical Finance*, 8, 291–323. [38]

Davidson, James (1994), *Stochastic Limit Theory*. Oxford University Press. [62]

DiCiccio, Cyrus J. and Joseph P. Romano (2017), "Robust permutation tests for correlation and regression coefficients." *Journal of the American Statistical Association*, 112, 1211–1220. [39]

Foster, Dean P. and Dan B. Nelson (1996), "Continuous record asymptotics for rolling sample variance estimators." *Econometrica*, 64, 139–174. [38]

Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw (2001), "Identification and estimation of treatment effects with a regression-discontinuity design." *Econometrica*, 69, 201–209. [38]

Hoeffding, Wassily (1952), "The large-sample power of tests based on permutations of observations." *Annals of Mathematical Statistics*, 3, 169–192. [45]

Horváth, Lagos and Gregory Rice (2014), "Extensions of some classical methods in change point analysis." *Test*, 23, 219–255. [39]

Hušková, Maria (2004), "Permutation principle and bootstrap in change point analysis." *Fields Institute Communications*, 44, 273–291. [39]

Imbens, Guido and Karthik Kalyanaraman (2011), "Optimal bandwidth choice for the regression discontinuity estimator." *The Review of Economic Studies*, 79, 933–959. [38]

Imbens, Guido W. and Thomas Lemieux (2008), "Regression discontinuity designs: A guide to practice." *Journal of Econometrics*, 142, 615–635. [38]

Jacod, Jean and Philip Protter (2012), *Discretization of Processes*. Springer. [38, 48, 49, 64]

Jentsch, Carsten and Markus Pauly (2015), "Testing equality of spectral densities using randomization techniques." *Bernoulli*, 21, 697–739. [39]

Kirch, Claudia (2007), "Block permutation principles for the change analysis of dependent data." *Journal of Statistical Planning and Inference*, 137, 2453–2474. [39]

Kirch, Claudia and Josef Steinebach (2006), "Permutation principles for the change analysis of stochastic processes under strong invariance." *Journal of Computational and Applied Mathematics*, 186, 68–88. [39]

Lee, David S. and Thomas Lemieux (2010), "Regression discontinuity designs in economics." *Journal of Economic Literature*, 48, 281–355. [38]

Lehmann, Erich L. and Joseph P. Romano (2005), *Testing Statistical Hypothesis*. Springer. [38, 42, 43, 45, 61, 64]

Li, Jia and Yunxiao Liu (2020), "Efficient estimation of integrated volatility functionals under general volatility dynamics." Econometric Theory. (Forthcoming). [48]

Li, Jia, Viktor Todorov, and George Tauchen (2017), "Jump regressions." *Econometrica*, 85, 173–195. [38]

Li, Jia, Viktor Todorov, and Qiushi Zhang (2021), "Testing the dimensionality of policy shocks." Review of Economics and Statistics. (Forthcoming). [59]

Li, Jia and Dacheng Xiu (2016), "Generalized method of integrated moments for high-frequency data." *Econometrica*, 84, 1613–1633. [40, 46]

Nakamura, Emi and Jón Steinsson (2018a), "High-frequency identification of monetary non-neutrality: The information effect." *Quarterly Journal of Economics*, 133, 1283–1330. [40, 59]

Nakamura, Emi and Jón Steinsson (2018b), "Identification in macroeconomics." *Journal of Economic Perspectives*, 32, 59–86. [46]

Stock, James (1994), "Unit Roots, Structural Breaks and Trends." In *Handbook of Econometrics*, Vol. 4 (R. Engle and D. McFadden, eds.), 2739–2841, Elsevier, Amsterdam. [38]

Thistlethwaite, Donald L. and Donald T. Campbell (1960), "Regression-discontinuity analysis: An alternative to the ex post facto experiment." *Journal of Educational Psychology*, 51, 309–317. [38]

Todorov, Viktor and George Tauchen (2012), "Realized Laplace transforms for pure-jump semimartingales." *The Annals of Statistics*, 40, 1233–1262. [41]

Vaart, Aad W. and Jon A. Wellner (1996), *Weak Convergence and Empirical Processes*. Springer-Verlag. [64]

Co-editor Andres Santos handled this manuscript.

Manuscript received 21 November, 2020; final version accepted 11 July, 2022; available online 22 July, 2022.