Weak Identification in News-Jump Regressions*

Jiezheng Wei†

Job Market Paper

October 29, 2026

(Updated regularly—latest version here)

Abstract

News-jump regressions are high-frequency event-study regressions under infill asymptotics. When news-jumps in explanatory variables over narrow observation windows are of the same magnitude as diffusive background noise, the regression coefficient is weakly identified, invalidating least-squares estimation and inference. I develop identification-robust inference based on an Anderson–Rubin test and derive, in a Gaussian limit experiment arising from the jump-diffusion process, a minimal sufficient and complete statistic for jumps, thereby enabling powerful conditional inference. In applications, I invert the proposed robust tests to revisit the empirical questions of Nakamura and Steinsson (2018) using minute-by-minute data, assess the effects of oil shocks triggered by the Russia–Ukraine and Iran wars, and examine the dependence of news-driven ETF jumps on market-index jumps.

Keywords: high-frequency event study regressions, infill asymptotics, news, jump regressions, weak identification, limit experiment.

*I am deeply indebted and grateful to my advisor, Jia Li, for his continuing support. My sincere thanks go to Dachuan Chen, Xu Cheng, Liyu Dou, Michal Kolesár, Yuichi Kitamura, Kenichi Nagasawa, Taisuke Otsu, Andrew Patton, Peter Phillips, Shuping Shi, Mikkel Sølvsten, Liyang Sun, Sen Zhang, Yichong Zhang, Liang Zhong and participants at the SMU Econometrics workshops for their helpful comments and discussions. All remaining errors are my own.

†School of Economics, Singapore Management University; E-mail: jzwei.2022@smu.edu.sg.

1 Introduction

News jumps, defined here as price jumps triggered by news and latent in observed changes around news announcements, reflect how financial markets incorporate information shocks into asset prices in an efficient market. Examining such high-frequency co-movements of related variables over narrow announcement windows is increasingly popular, as this instantaneous market reaction enables empiricists to tease out causal effects through exogenous variation generated by quasi-experiments in a discontinuity-based identification scheme; see Gürkaynak, Swanson, and Wright (2026).

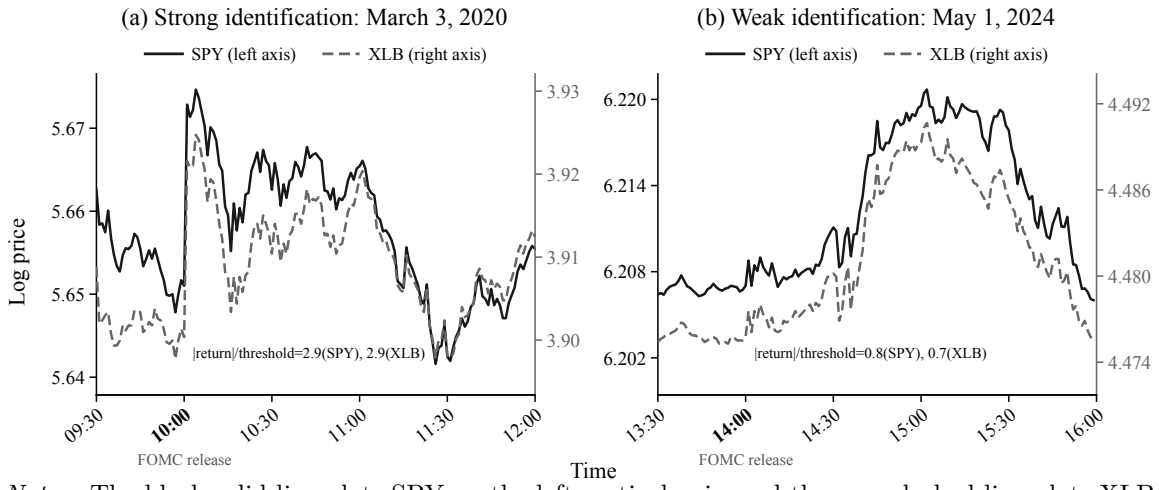
Despite the prevalence of high-frequency event study regressions using 30-minute or wider windows under long-span asymptotics in the high-frequency identification literature[1], the shock-induced co-jump nature can be studied using finer, e.g., 1-minute, data under infill asymptotics. Li, Todorov, and Tauchen (2017, henceforth, LTT) develop jump regressions in a statistically valid two-step procedure. For two related high-frequency time series, the first step detects jumps in the right-hand-side variable using the standard truncation method[2]; the second step runs least-squares regressions on the filtered, large and rare jumps, and conducts inference on the regression coefficient, known as the *jump beta*. For high-frequency event studies, this infill-asymptotic approach offers three benefits: (i) minimizing omitted-variable concerns that confound identification; (ii) avoiding the need for long-span samples; (iii) being able to capture the time-varying feature of causal parameters.

For high-frequency event studies, announcement times help localize news jumps, which provides a clear economic interpretation, allowing researchers to bypass first-stage truncation for jump detection. However, high-frequency changes in explanatory variables around news announcements are often so small that they fall below the truncation threshold. Such small variation reflects either no jump or a genuine but undetected small jump given the finite sample discretization. Consequently, regression analysis based on these small jumps risks size-distorted inference and misleading results. This "small-shock" issue is well recognized in empirical macroeconomics, notably as the "power problem" emphasized by Nakamura and Steinsson (2018).

[1] See Nakamura and Steinsson (2018), Känzig (2021) for representative work on monetary policy and oil supply news shocks, respectively, and the survey by Ramey (2016).

[2] In high-frequency financial econometrics, the truncation method detects jumps by choosing a threshold shrinking more slowly than the diffusion so as to discriminate diffusive components from jumps under infill asymptotics; see Mancini (2001) and Chapter 9 of Jacod and Protter (2012).

Figure 1: One-minute log prices of SPY and XLB around FOMC announcements



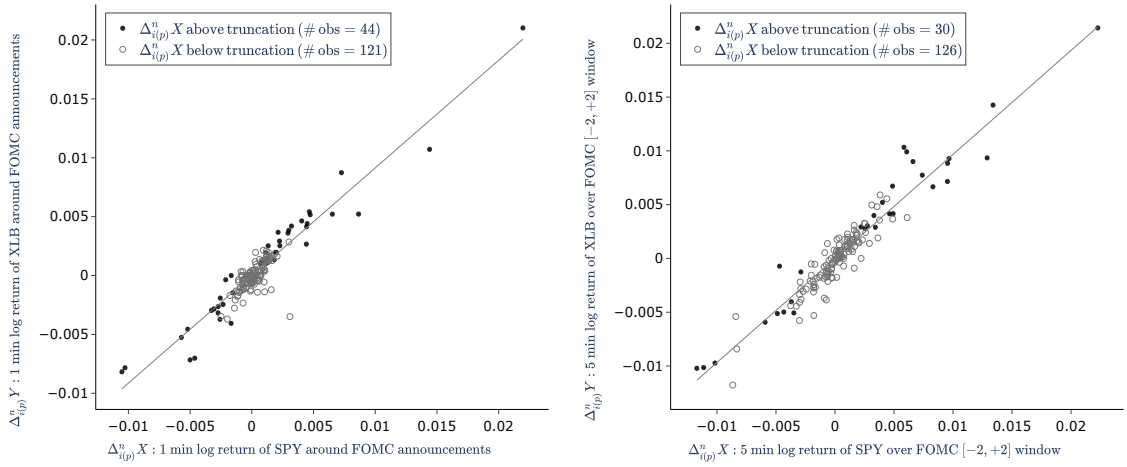
Notes: The black solid line plots SPY on the left vertical axis, and the gray dashed line plots XLB on the right. The in-panel quantity is the absolute release-minute log return divided by the asset-specific truncation threshold from equation (25), with truncation exponent 0.49. It exceeds one for both assets in Panel (a) but is below one for both assets in Panel (b).

To illustrate the small-jump problem, I examine SPY and XLB log returns around Federal Open Market Committee (FOMC) announcements. Figure 1 contrasts two cases: the unscheduled 50-basis-point rate cut on March 3, 2020, when both one-minute returns around announcements exceeded the truncation threshold, and the anticipated no-change decision on May 1, 2024, when neither did. Figure 2 shows that weak first stages are pervasive: 121 of 165 release-minute return pairs fall below the thresholds over 2003–2024, and 126 of 156 remain below after aggregation over the five-minute FOMC window $[t-2, t+2]$. Thus, even major macro-announcements often generate price responses too small to be statistically identified as jumps. This motivates the development of inference robust to small or absent jumps.

I formulate the small-jump phenomenon as a weak identification problem due to its low signal-to-noise ratio[3]. I model jump magnitudes as shrinking at the inverse square-root rate of the sample size and, by analogy with the weak-instrument literature (Staiger and Stock (1997)), refer to such jumps as *weak jumps*. The theory is designed for moderately high-frequency data such as one-minute or coarser sampling, where market microstructure effects are less pronounced. The square-root shrinkage rate

[3]Here, "noise" is used in the statistical sense, rather than market microstructure noise. Under the continuous-time model, each log return comprises drift, diffusion, and jump components, with diffusion constituting the dominant noise for the jump signal, as drift is asymptotically negligible at high frequencies.

Figure 2: The scatter plot of high-frequency return pairs around FOMC announcements.



Notes: The panels plot XLB against SPY log returns around FOMC announcements, 2003–2024, using one-minute and five-minute $[t-2, t+2]$ windows. The right panel has 9 fewer observations owing to missing early XLB data. Lines are zero-intercept fits of observations above truncation.

places weak jumps on the diffusion scale, echoing the common practice of treating the small-jump component of Lévy processes as Gaussian noise (Carpentier, Duval, and Mariucci (2021)).

Writing the two systems in parallel clarifies the connection between news-jump regressions and instrumental variables (IV) regressions:

$$\begin{cases} \Delta^n_{i(p)}X = J^X_p + \varepsilon^X_{i(p)}, \\ \Delta^n_{i(p)}Y = J^X_p\beta + \varepsilon^Y_{i(p)}. \end{cases} \qquad \begin{cases} X_i = Z_i\pi + V_{i1}, \\ Y_i = Z_i\pi\beta + V_{i2}. \end{cases} \tag{1}$$

(News-jump regression) *(Reduced form IV regression)*

In the news-jump regression on the left, $\left(\Delta^n_{i(p)}X, \Delta^n_{i(p)}Y\right)$ denotes the return pair over the interval containing the pth announcement, J^X_p the news-induced X-jump, and $\left(\varepsilon^X_{i(p)}, \varepsilon^Y_{i(p)}\right)$ the continuous return components; in the IV regression on the right, X_i is the single endogenous regressor, Z_i its instrument, and (V_{i1}, V_{i2}) the reduced-form disturbances. Hence, J^X_p and $Z_i\pi$ play analogous first-stage roles: each enters X directly and Y through β; when either J^X_p or π is local to zero, identification of β becomes weak. Nevertheless, this setting differs from weak IV regressions in three respects. First, the jump J^X_p is latent while Z_i is observed. Second, the infill asymptotics considered here refine the sampling grid around a fixed set of announcements,

whereas conventional weak-IV and event-study asymptotics increase the number of events in an outfill (long-span) manner. Third, the announcement-specific covariance of the continuous components $\left(\varepsilon^X_{i(p)}, \varepsilon^Y_{i(p)}\right)$ is locally estimable from neighboring high-frequency returns, whereas conventional IV inference handles reduced-form dependence through assumptions such as i.i.d. errors or a Kronecker-product covariance structure, or through HAC estimation of the long-run variance (Moreira and Moreira (2019)).

This paper makes three contributions. First, I show in theory and simulations that under weak jumps, LTT's least-squares regression estimation is no longer consistent and regression-based tests are size-distorted. Second, I provide an identification-robust and asymptotically similar Anderson–Rubin test (Anderson and Rubin (1949)) allowing for price-volatility co-jumps[4]. Third, under the assumption that news announcement times are observed and equal to jump times, I find a sufficient statistic for jumps in the Gaussian limit experiment arising from jump-diffusion modelling to develop conditional inference, as in Moreira (2003) and Andrews (2016a). Overall, this paper provides constructive econometric solutions for weakly identified continuous-time event study models with high-frequency data.

This paper relates to the literature on jump dependence, weak identification, and high-frequency identification. Todorov and Bollerslev (2010) propose that jumps should be priced differently from continuous risk, thereby introducing jump beta. Li, Todorov, and Tauchen (2017) establish regression analysis for jumps, including a specification test, $\sqrt{n}$-consistency, and a semiparametric efficiency result. The close interrelation between news and jumps/co-jumps has been well documented; see Anderson, Bollerslev, Diebold, and Vega (2003), Maheu and McCurdy (2004), Lahaye, Laurent, and Neely (2011), Christensen, Timmermann, and Veliyev (2025) among many others. However, despite the extensive findings, few studies have been devoted to small jumps surrounding news, which presents its own challenges for identification and inference.

As jump regressions bear a high-level resemblance to regression discontinuity (Bollerslev, Li, and Xue (2018)) and, as revealed by (1), to IV regression, the theory developed here naturally parallels the weak identification literature developed in

[4]Unlike in weak-IV regressions, price-volatility co-jumps are an additional complication in high-frequency data: the diffusive noise level accompanying the price-jump signal may jump simultaneously; see Jacod and Todorov (2010) and LTT.

these settings; see Anderson and Rubin (1949), Staiger and Stock (1997), Kleibergen (2002, 2005), Moreira (2003), Andrews, Moreira, and Stock (2006), Feir, Lemieux, and Marmer (2016), Andrews (2016a). In the presence of microstructure noise, the notion of weakly identified jumps is adopted by Li (2013).

The empirical literature on high-frequency identification is closely related to the applied purpose of this paper; see Kuttner (2001), Gürkaynak, Sack, and Swanson (2005), Nakamura and Steinsson (2018), Bauer and Swanson (2023b), among others. Anderson, Bollerslev, Diebold, and Vega (2003) show that the evidence is weaker using daily than intradaily data. These studies typically adopt discrete-time asymptotics within ordinary least squares (OLS) regressions or vector autoregressions (VARs). Stock and Watson (2012) refer to high-frequency monetary shocks as "external instruments" that facilitate identification in structural VARs and find a weak-instrument problem. Casini and McCloskey (2026) have similar findings that "narrow size of the window is not sufficient for identification" in high-frequency event-study regressions, a concern that this paper addresses as weak identification.

The rest of the paper is organized as follows. Section 2 introduces the setup and Section 3 studies weak-jump asymptotics. Sections 4 and 5 develop identification-robust inferential methods. Sections 6 and 7 report simulation and empirical results, respectively. Section 8 concludes. Testing procedures and proofs of the main theorems are presented in the appendices.

The following notation is used throughout. Let $\mathbb{R}$ denote the real line, $\mathbb{R}_* \equiv \mathbb{R} \setminus \{0\}$, $\mathbb{R}_+ \equiv [0, \infty)$ and $\mathbb{R}_{++} \equiv (0, \infty)$. For a matrix A, $A^\intercal$ denotes its transpose. Write $x_n \asymp y_n$ if $x_n/c \leq y_n \leq cx_n$ for some constant $c \geq 1$ and all n. Use $\|\cdot\|$ for the Euclidean norm or matrix operator norm, $\mathcal{M}_2$ for the set of 2×2 positive semidefinite matrices, $|\mathcal{P}|$ for the cardinality of a set $\mathcal{P}$, and $\perp\!\!\!\perp$ for independence. Finally, $\xrightarrow{\mathcal{L}}$, $\xrightarrow{\mathcal{L}\text{-}s}$, and $\xrightarrow{\mathbb{P}}$ denote convergence in law, stable convergence in law, and convergence in probability, respectively.

2 Setup and Assumptions

Consider a bivariate process $Z_t \equiv (X_t, Y_t)^\intercal$, such as log prices or interest rates, defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$. Assume that Z_t admits the following

decomposition into continuous and jump components:

$$Z_t = Z_0 + \int_0^t b_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + \sum_{s \le t} \Delta Z_s. \tag{2}$$

Here, the drift process $b \in \mathbb{R}^2$ and the stochastic volatility process $\sigma \in \mathcal{M}_2$ are both càdlàg adapted, and $c_t = \sigma_t \sigma_t^\intercal$ denotes the spot covariance matrix. The driving process $W = (W_X, W_Y)^\intercal$ is a standard two-dimensional Brownian motion, and $\Delta Z_t \equiv Z_t - Z_{t-}$ is a pure jump process.

Suppose an economist is interested in a type of news, labeled by $\mathfrak{a}$. Let $\mathfrak{a}_p$ denote the timestamp of its pth announcement, which may be from either prescheduled or random events. Let τ_p denote the ocurring time of the X-jump triggered by the pth announcement. Meanwhile, the process $(Z_t)_{t\in[0,T]}$ is discretely observed at $t = 0, \Delta_n, \ldots, n\Delta_n$, where Δ_n denotes the time mesh between consecutive observations, and the sample size $n = \lfloor T/\Delta_n \rfloor$. The asymptotic scheme is infill, i.e., $\Delta_n \to 0$ or equivalently $n \to \infty$. The i-th increment is obtained by a difference operator Δ_i^n defined as follows:

$$\Delta_i^n Z \equiv Z_{i\Delta_n} - Z_{(i-1)\Delta_n}.$$

Let the $i(p)$th sampling interval contain the pth announcement time $\mathfrak{a}_p$. Additionally, assume τ_p is in the same interval, i.e., $\tau_p \in ((i(p)-1)\Delta_n, i(p)\Delta_n]$. Define the relative location of the pth news-jump within its sampling interval as

$$\kappa_{n,p} \equiv \frac{\tau_p - (i(p)-1)\Delta_n}{\Delta_n} \in (0, 1] \tag{3}$$

The assumption about locating news-jumps by news-announcement times is formalized as follows.

Assumption 1. For each Δ_n and each $p \in \mathcal{P}_{\mathfrak{a}_B}$, the location index $i(p)$ for the pth announcement is observed, and the news-jump time $\tau_p \in [(i(p)-1)\Delta_n, i(p)\Delta_n)$. Moreover, $(\kappa_{n,p})_p \xrightarrow{\mathcal{L}\text{-}s} (\kappa_p)_p$ where κ_p's are i.i.d. uniformly over $(0, 1]$.

Assumption 1 allows some ambiguity about the exact timing of the news event, but requires the associated news jump to be localized within the Δ_n-discretization and $\tau_p = \mathfrak{a}_p + o_{\mathbb{P}}(\Delta_n)$, so that information leakage or delayed market adjustment is asymptotically negligible. In practice, if the price response begins before the $i(p)$th

interval or unfolds over several intervals, especially for large jumps, the event interval should be re-localized accordingly. Finally, imposing the uniformly distribution of within-interval jump time location reflects the property of Poisson-type jumps and accords with Theorem 4.3.1 and Theorem 4.4.10 Jacod and Protter (2012) in the setting of Itô semimartingales.

Let B be a Borel subset of the time domain $\mathbb{R}_+$ with finite Lebesgue measure. Define $\mathfrak{a}_B \equiv \{\mathfrak{a}_p : \mathfrak{a}_p \in B\}$ as the collection of all announcement times in B, and let $\mathcal{P}_{\mathfrak{a}_B} = \{p \geq 1 : \mathfrak{a}_p \in \mathfrak{a}_B\}$ denote the ordinal indexes of news jumps related to $\mathfrak{a}_B$. Assume a linear relationship between the jumps, motivated by the empirical large-jump pattern shown in Figure 2, namely

$$\Delta Y_{\tau_p} = \beta \Delta X_{\tau_p}, \quad \text{for } p \in \mathcal{P}_{\mathfrak{a}_B}, \tag{4}$$

where β is the so-called jump beta. Moreover, equation (4) implies the identification: whenever $\Delta X_{\tau_p} \neq 0$, the ratio $\Delta Y_{\tau_p}/\Delta X_{\tau_p}$ equals β. The testing problem is

$$H_0 : \beta = \beta_0 \quad \text{against} \quad H_1 : \beta \neq \beta_0, \tag{5}$$

where $\beta_0 \in \mathbb{R}$ is picked by the economist. Combining (2) and (4), the news-jump regression is

$$\begin{aligned} &\Delta^n_{i(p)} Y = \beta \Delta^n_{i(p)} X + u_{p,n}(\beta), \quad \text{where} \\ &u_{p,n}(\beta) = (-\beta, 1) \left(\int_{(i(p)-1)\Delta_n}^{i(p)\Delta_n} b_s \, \mathrm{d}s + \int_{(i(p)-1)\Delta_n}^{i(p)\Delta_n} \sigma_s \, \mathrm{d}W_s \right), \quad \text{for } p \in \mathcal{P}_{\mathfrak{a}_B} \end{aligned}$$

and the error term satisfies $u_{p,n}(\beta) = O_{\mathbb{P}}(\Delta_n^{1/2})$ under regularity conditions specified later in Assumption 3.

To account for small jumps that fall below the standard truncation threshold, impose the following asymptotic structure on jump magnitudes:

$$\Delta X_{\tau_p} = O_{\mathbb{P}}(\Delta_n^r) = O_{\mathbb{P}}(n^{-r}), \quad r \geq 0.$$

The shrinkage rate r governs the strength of identification. When $r = 0$, the model corresponds to the conventional strongly identified case studied in LTT; when $0 < r < 1/2$, identification is semi-strong; and when $r = 1/2$, identification is weak; see Andrews and Cheng (2012). The analysis focuses on the weakly identified case, i.e.

$r = 1/2$.

Assumption 2 (Weak jumps)**.** For each announcement in $\mathfrak{a}_B$, the news jump of the X-process is *weak* in the sense that

$$\Delta X_{\tau_p} = \xi_p \Delta_n^{1/2}, \quad p \in \mathcal{P}_{\mathfrak{a}_B},$$

where ξ_p is an $\mathbb{R}$-valued $\mathcal{F}_{\tau_p}$-measurable random variable for each $p \in \mathcal{P}_{\mathfrak{a}_B}$.

Two remarks on Assumption 2 are in order. First, the exponent of one-half places jumps on the same scale as diffusion. Second, when treating jumps as nuisance parameters, the local specification of weak jumps should be interpreted as an asymptotic device for capturing a setting in which increasing the sampling frequency does not help identification, rather than as a literal assertion that jump sizes decrease with the sample size. This specification delineates the weak-identification regime in which conventional Wald inference will fail, whereas identification-robust inference should remain valid, see Dufour (1997). The following regularity conditions are also imposed.

Assumption 3 (Regularity conditions)**.** (a) The drift process b_t is locally bounded and adapted; (b) the volatility process σ_t is càdlàg adapted and positive semi-definite for any $t \in [0, T]$; (c) no jump occurs on the boundary of B, almost surely; (d) the news-jump process $\sum_{p \geq 1} \Delta Z_{\tau_p} 1_{\{\tau_p \leq t\}}$ is finitely active.

Assumption 3(a) renders the drift component negligible relative to the diffusion part; Assumption 3(b) allows the existence of price-volatility co-jumps; Assumption 3(c) can be easily satisfied by slightly enlarging the considered time region B; Assumption 3(d) is justified by the focus on finitely many scheduled news announcements.

3 Weak-Jump Asymptotics

3.1 A Review of the Strong Identification Case

This subsection briefly reviews LTT's jump regression under strong identification. Define

$$R_p \equiv \sqrt{\kappa_p}\, \sigma_{\mathfrak{a}_p -} \Psi_{p-} + \sqrt{1 - \kappa_p}\, \sigma_{\mathfrak{a}_p} \Psi_{p+}, \quad \varsigma_p(\beta_0) \equiv (-\beta_0, 1) R_p, \tag{6}$$

where $(\kappa_p, \Psi_{p-}, \Psi_{p+})_{p\geq 1}$ is a sequence of triplets of mutually independent random variables, with $(\kappa_p)_{p\geq 1}$ being i.i.d. uniformly distributed on $[0,1)$, and both $(\Psi_{p-})_{p\geq 1}$ and $(\Psi_{p+})_{p\geq 1}$ being i.i.d. bivariate Gaussian vectors with identity covariance matrices. Moreover, consistent estimators for the spot pre-jump and post-jump volatilities are needed. Set

$$k_n = \lfloor c\Delta_n^{-\gamma} \rfloor, \quad \text{for some constant } c \text{ and } 0 < \gamma < 1,$$

where $k_n \to \infty$ ensures consistency, and $k_n\Delta_n \to 0$ guarantees the estimation targets the spot volatility. At each i-th observation, the large-k estimators are

$$\begin{cases} \hat{c}_{n,i+} = \dfrac{\sum_{j=1}^{k_n} \Delta_{i+j}^n Z \Delta_{i+j}^n Z^\intercal 1_{\{-u_n \leq \Delta_{i+j}^n Z \leq u_n\}}}{\Delta_n \sum_{j=1}^{k_n} 1_{\{-u_n \leq \Delta_{i+j}^n Z \leq u_n\}}}, \\ \hat{c}_{n,i-} = \dfrac{\sum_{j=1}^{k_n} \Delta_{i-j}^n Z \Delta_{i-j}^n Z^\intercal 1_{\{-u_n \leq \Delta_{i-j}^n Z \leq u_n\}}}{\Delta_n \sum_{j=1}^{k_n} 1_{\{-u_n \leq \Delta_{i-j}^n Z \leq u_n\}}} \end{cases} \tag{7}$$

where $u_n = (u_n^X, u_n^Y)^\intercal \in \mathbb{R}^2_{++}$ is the truncation threshold, with $u_n^X \asymp \Delta_n^\varpi$ and $u_n^Y \asymp \Delta_n^\varpi$ for $\varpi \in (0, 1/2)$. Based on Section 3.2 of LTT, the OLS estimator is

$$\tilde{\beta}_n = \tilde{\beta}_n(\mathfrak{a}_B) = \frac{\sum_{i\in\mathcal{I}(\mathfrak{a}_B)} \Delta_i^n X \Delta_i^n Y}{\sum_{i\in\mathcal{I}(\mathfrak{a}_B)} \Delta_i^n X^2}. \tag{8}$$

Let the weighting function $w : \mathcal{M}_2 \times \mathcal{M}_2 \times \mathbb{R} \to \mathbb{R}$ be a continuous function. The weighted least squares (WLS) estimator is

$$\hat{\beta}_n(\mathfrak{a}_B, w) = \frac{\sum_{i\in\mathcal{I}(\mathfrak{a}_B)} w(\hat{c}_{n,i-}, \hat{c}_{n,i+}, \tilde{\beta}_n) \Delta_i^n X \Delta_i^n Y}{\sum_{i\in\mathcal{I}(\mathfrak{a}_B)} w(\hat{c}_{n,i-}, \hat{c}_{n,i+}, \tilde{\beta}_n) \Delta_i^n X^2}. \tag{9}$$

Moreover, the optimal weight corresponding to the generalized least squares (GLS) estimator is $w^*(\hat{c}_{n,i-}, \hat{c}_{n,i+}, \tilde{\beta}_n) = 2/(-\tilde{\beta}_n, 1)(\hat{c}_{n,i-} + \hat{c}_{n,i+})(-\tilde{\beta}_n, 1)^\intercal$ by accounting for heteroskedasticity. LTT's result implies that, under Assumptions 1 and 3,

$$\Delta_n^{-1/2}(\hat{\beta}_n(\mathfrak{a}_B, w) - \beta) \xrightarrow{\mathcal{L}\text{-}s} \frac{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} w(c_{\mathfrak{a}_p-}, c_{\mathfrak{a}_p}, \beta) \Delta X_{\mathfrak{a}_p} \varsigma_p}{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} w(c_{\mathfrak{a}_p-}, c_{\mathfrak{a}_p}, \beta) \Delta X_{\mathfrak{a}_p}^2}. \tag{10}$$

The above result clearly subsumes the $\sqrt{n}$-consistency of OLS and GLS by respectively letting $w = 1$ and $w = w^*$.

3.2 Jump Regressions under Weak-Jump Asymptotics

Given the weak-jump setup, I re-examine LTT's jump regressions. Define $\mathcal{I}(\mathfrak{a}_B) = \{i(p) : p \in \mathcal{P}_{\mathfrak{a}_B}\}$ to collect all the sampling intervals in B that contain news announcements $\mathfrak{a}$. In the first step of LTT, the truncation method on $\mathcal{I}(\mathfrak{a}_B)$ generates:

$$\mathcal{I}_n(\mathfrak{a}_B) = \{i \in \mathcal{I}(\mathfrak{a}_B) : |\Delta_i^n X| > u_n\}, \quad u_n \asymp \Delta_n^{\varpi}, \quad \varpi \in (0, 1/2). \tag{11}$$

Clearly, $\mathcal{I}_n(\mathfrak{a}_B) \subseteq \mathcal{I}(\mathfrak{a}_B)$. The asymptotic results of news-jump regression under weak jumps are summarized in the following theorem.

Theorem 1 (Weak-jump asymptotics)**.** Under Assumptions 1, 2, and 3, the following results hold:

(a) $\mathbb{P}(\mathcal{I}_n(\mathfrak{a}_B) = \varnothing) \to 1$ as $n \to \infty$.

(b) For the ordinary least squares estimator defined by (8), the limiting distribution is

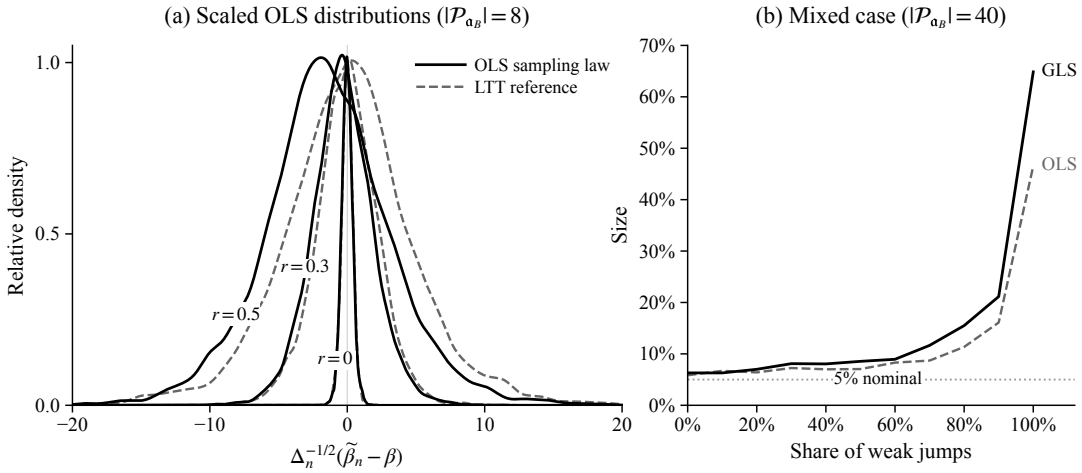
$$\tilde{\beta}_n - \beta \xrightarrow{\mathcal{L}\text{-}s} \tilde{b} \equiv \frac{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} (e_1^\intercal R_p + \xi_p)\varsigma_p(\beta)}{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} (e_1^\intercal R_p + \xi_p)^2} = O_{\mathbb{P}}(1), \tag{12}$$

where $e_1 = (1, 0)^\intercal \in \mathbb{R}^2$, and R_p and $\varsigma_p(\beta)$ are defined in (6). For the weighted least squares estimator defined by (9), the limiting distribution is

$$\hat{\beta}_n(\mathfrak{a}_B, w) - \beta \xrightarrow{\mathcal{L}\text{-}s} b_w = \frac{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} w(c_{\mathfrak{a}_p -}, c_{\mathfrak{a}_p}, \beta + \tilde{b}) \left(e_1^\intercal R_p + \xi_p\right) \varsigma_p(\beta)}{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} w(c_{\mathfrak{a}_p -}, c_{\mathfrak{a}_p}, \beta + \tilde{b}) \left(e_1^\intercal R_p + \xi_p\right)^2}$$

Theorem 1(a) indicates that the standard truncation method asymptotically fails to detect any weak jumps, which is aligned with the empirical concern that small jumps are undetectable. Theorem 1 (b) shows that the WLS estimators $\hat{\beta}_n(\mathfrak{a}_B, w)$, which include OLS and GLS as special cases, are inconsistent under weak jumps and hard to bias-correct, reflecting a typical consequence of weak identification. Consequently, weak-jump limits are substantially heavier-tailed than the strong-identification limits, e.g., in the just-identified case, the limit (10) is Gaussian, whereas (12) is a nonstandard fat-tailed random variable in ratio form even without scaling. The size distortion is evident from the simulation in Figure 3: its left panel shows that the OLS limit increasingly departs from its strongly identified counterpart as jumps

Figure 3: Scaled OLS law under weak jumps versus the LTT benchmark



Notes: Panel (a) compares the simulated sampling law of $\Delta_n^{-1/2}(\tilde{\beta}_n - \beta)$ (solid) with the limiting distribution of LTT (dashed) for shrinkage rate $r \in \{0, 0.3, 0.5\}$ when $\Delta_n = 1/390$. Each pair is normalized by its LTT density at zero; densities are kernel-smoothed. Panel (b) reports null rejection rates of LTT-OLS/GLS as the share of weak jumps rises; $|\mathcal{P}_{\mathfrak{a}_B}|$ denotes the number of jumps.

shrink, and the right panel shows the size distortion in a mixed case: while a small portion of weak jumps is negligible, the size of the LTT procedures rises sharply above the nominal level once weak jumps dominate the sample. These issues necessitate the development of identification-robust inference.

4 The Anderson–Rubin Test

In this section, I develop the Anderson–Rubin test which is valid for any jump size by invoking the null hypothesis. First consider the just-identified case of a single announcement. Recalling the testing problem (5), I propose the following Anderson–Rubin (AR) statistic on the pth announcement:

$$S_{n,p}(\beta_0) = \Delta_n^{-1/2} \left(\Delta_{i(p)}^n Y - \beta_0 \Delta_{i(p)}^n X \right), \quad p \in \mathcal{P}_{\mathfrak{a}_B}. \tag{13}$$

The AR statistic (13) exploits the specification (4) to cancel jumps. If β_0 coincides with the true parameter β, then the X-jump and Y-jump components cancel in $S_{n,p}(\beta_0)$ by the linear jump relation, yielding a null distribution free of the jump size.

Theorem 2 establishes that under the null, $S_{n,p}(\beta_0)$ will stably converge in law to

$\varsigma_p(\beta_0)$ defined in (6). I then aggregate $\{S_{n,p}(\beta_0)\}_{p\in\mathcal{P}_{\mathfrak{a}_B}}$ in L_2 norms and use simulated critical values, as detailed in Algorithm 1.

Theorem 2. Restricting attention to $\Omega(\mathfrak{a}_B)$, suppose Assumptions 1, 2, and 3 are satisfied. Then:

(a) The statistic satisfies $S_{n,p}(\beta_0) \xrightarrow{\mathcal{L}\text{-}s} \varsigma_p(\beta_0)+(\beta-\beta_0)\xi_p$ where $\varsigma_p(\beta_0) = (-\beta_0, 1)R_p$ and $R_p = \sqrt{\kappa_p}\sigma_{\mathfrak{a}_p-}\Psi_{p-} + \sqrt{1-\kappa_p}\sigma_{\mathfrak{a}_p}\Psi_{p+}$ for an independent uniform random variable κ_p over the unit interval $[0, 1)$ and independent bivariate normals Ψ_{p-} and Ψ_{p+}.

(b) For $\mathfrak{a}_B$, define $Q_S \equiv \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} [S_{n,p}(\beta_0)]^2/\hat{v}_{S,p}(\beta_0)$ where $\hat{v}_{S,p}(\beta_0) \equiv \frac{1}{2}h_0(\hat{c}_{n,i(p)-} + \hat{c}_{n,i(p)+})h_0^\intercal \xrightarrow{\mathbb{P}} v_{S,p}(\beta_0) \equiv \frac{1}{2}h_0(c_{\mathfrak{a}_p-} + c_{\mathfrak{a}_p})h_0^\intercal$ which is the conditional variance of $\varsigma_p(\beta_0)$ given $(c_{\mathfrak{a}_p-}, c_{\mathfrak{a}_p})$, and $h_0 = (-\beta_0, 1)^\intercal$. Under the null $H_0 : \beta = \beta_0$,

$$Q_S \xrightarrow{\mathcal{L}\text{-}s} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{[\varsigma_p(\beta_0)]^2}{v_{S,p}(\beta_0)}. \tag{14}$$

(c) The test based on (14) has asymptotic size $\alpha \in (0, 1)$ for every $\beta_0 \in \mathbb{R}$, in the sense that $\lim_{n\to\infty} \mathbb{P}\left(Q_S \geq cv^\alpha_{n,AR}(\beta_0) \mid H_0\right) = \alpha$ for every $\beta_0 \in \mathbb{R}$, and $cv^\alpha_{n,AR}(\beta_0)$ is calculated by the procedure described in Algorithm 1.

Algorithm 1. (i) Generate one realization from a sequence of random triplets $(\tilde{\kappa}_p, \tilde{\Psi}_{p-}, \tilde{\Psi}_{p+})_{p\in\mathcal{P}_{\mathfrak{a}_B}}$ which is an independent copy of $(\kappa_p, \Psi_{p-}, \Psi_{p+})_{p\in\mathcal{P}_{\mathfrak{a}_B}}$, and calculate for each p

$$\tilde{R}_{n,p} = \sqrt{\tilde{\kappa}_p}\hat{\sigma}_{n,i(p)-}\tilde{\Psi}_{p-} + \sqrt{1-\tilde{\kappa}_p}\hat{\sigma}_{n,i(p)+}\tilde{\Psi}_{p+}, \quad \tilde{\varsigma}_{n,p}(\beta_0) = (-\beta_0, 1)\tilde{R}_{n,p}$$
$$\hat{v}_{S,p}(\beta_0) = \frac{1}{2}(-\beta_0, 1)(\hat{c}_{n,i(p)-} + \hat{c}_{n,i(p)+})(-\beta_0, 1)^\intercal$$

where $(\hat{\sigma}_{n,i(p)-}, \hat{\sigma}_{n,i(p)+})$ are Cholesky decompositions of $\left(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}\right)$ defined in (7). Then store the value $MC = \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{[\tilde{\varsigma}_{n,p}(\beta_0)]^2}{\hat{v}_{S,p}(\beta_0)}$ in this one trial.

(ii) Given a significance level $\alpha \in (0, 1)$, independently repeat step (i) a large number of times and obtain a Monte Carlo sample of MC. Set $cv^\alpha_{n,AR}(\beta_0)$ as the $(1-\alpha)$-quantile of MC in the Monte Carlo sample.

By Theorem 2(a), the null distribution of the AR test at the pth announcement is a Gaussian mixture. The random variable κ_p locating the jump within the $i(p)$-th sampling interval is uniformly distributed according to the Itô semimartingale model. Under the null, Q_S is free of jumps and therefore robust to weak identification. Under the alternative, $(\beta - \beta_0)\xi_p$ provides power for the test and the power increases as $|\beta - \beta_0|$ grows. Under strong or semi-strong identification, the AR test is consistent against a fixed alternative since $\xi_p = O_{\mathbb{P}}(n^{1/2-r})$ and diverges as n increases for $r \in [0, 1/2)$. In Theorem 2(b), the variance normalization in Q_S places the announcement-specific AR tests on a common scale; without it, high-volatility announcements would mechanically receive greater weights. Theorem 2(c) permits test inversion for β. Accordingly, define the AR test $\varphi^{\alpha}_{AR} \equiv 1\{Q_S > cv^{\alpha}_{n,AR}(\beta_0)\}$ in which 1 means rejection. The identification-robust confidence set (CS) at confidence level $1 - \alpha$ is

$$\mathrm{CS}^{1-\alpha}_{AR} \equiv \{\beta_0 \in \mathbb{R} : \varphi^{\alpha}_{AR}(\beta_0) = 0\}.$$

It should be noted that $\mathrm{CS}^{1-\alpha}_{AR}$ does not spuriously recover the beta for continuous return as it is a nuisance parameter that is the same under the null and the alternative and does not contribute to distributional discrepancy for hypothesis testing.

5 Inference in the Limit Experiment

Although the AR test is robust to identification strength and can be shown to be optimal under just identification, it is inefficient under strong identification in overidentified settings ($|\mathcal{P}_{\mathfrak{a}_B}| \geq 2$); see Kleibergen (2005), Andrews (2016a). This section develops conditional inference (Moreira (2003), Chapter 10 of Lehmann and Romano (2005)) with the aim of deriving powerful tests.

To simplify the analysis, I directly consider testing in the limit problem within Le Cam's framework (Müller (2011), Andrews (2016a), Bollerslev, Li, and Ren (2024)) and derive an asymptotic result in the Supplement, showing that the related statistics in the discretized sample stably converge in law to the statistics in the limit problem, to formally justify this limit-of-experiment approach. The analysis in this section is based on Assumption 1′.

5.1 Conditional Inference in the Gaussian Limit Problem

The following assumption simplifies the limit experiment from a Gaussian mixture to a Gaussian shift experiment in which a sufficient statistic for jumps can be found due to the availability of a closed-form likelihood.

Assumption 1′. For every $p \in \mathcal{P}_{\mathfrak{a}_B}$, the realized announcement time $\mathfrak{a}_p$ is observed, and $\tau_p = \mathfrak{a}_p$ almost surely, and $\kappa_{n,p} = [\mathfrak{a}_p - (i(p)-1)\Delta_n]/\Delta_n \to \bar{\kappa}_p$ as $n \to \infty$ for some constant $\bar{\kappa}_p \in [0,1]$.

Assumption 1′ is plausible in empirically relevant settings. For example, pre-scheduled FOMC announcements are timestamped to the second, and SPY is highly liquid. As in Assumption 1, information leakage and market reaction time is simplified to be negligible, and Assumption 1′ strengthens this requirement by imposing $\tau_p = \mathfrak{a}_p$ with probability one.

This assumption is for theoretical simplification to derive the Gaussian shift limit experiment. One can avoid this assumption by standardizing $\Delta^n_{i(p)}Z$ and using a more involved analysis. What is more, Assumption 1′ is reasonable for pre-scheduled news releases in the sense that for FOMC announcements, one can set $\bar{\kappa}_p = 0$.

Define $g = (1,\beta)^\intercal$ and $h = (-\beta,1)^\intercal$ for any $\beta \in \mathbb{R}$, and let g_0 and h_0 denote their counterparts with β replaced by β_0. Let $\delta \equiv \beta - \beta_0$ denote the deviation from the null, so that $g = g_0 + \delta e_2$ where $e_2 = (0,1)^\intercal$. Under Assumptions 1′, 2, 3, in restriction to $\Omega(\mathfrak{a}_B)$, the limit experiment for $\Delta_n^{-1/2}\left(\Delta^n_{i(p)}X, \Delta^n_{i(p)}Y\right)^\intercal, p \in \mathcal{P}_{\mathfrak{a}_B}$ is Gaussian, namely

$$\frac{\Delta^n_{i(p)}Z}{\Delta_n^{1/2}} \xrightarrow{\mathcal{L}\text{-}s} Z_p \equiv (X_p, Y_p)^\intercal \sim \mathcal{N}\left(\xi_p g, \Sigma_p\right), \quad \Sigma_p \equiv \bar{\kappa}_p c_{\mathfrak{a}_p-} + (1-\bar{\kappa}_p)c_{\mathfrak{a}_p} \tag{15}$$

and Z_p's are independent across $p \in \mathcal{P}_{\mathfrak{a}_B}$. Here, the rescaled jump ξ_p is the nuisance parameter and β in g is the parameter of interest. Next, define two statistics exploiting the null candidate β_0. The first one is a sufficient statistic for the rescaled jump ξ_p, which captures the strength of the jump signal:

$$T_p \equiv \frac{g_0^\intercal \Sigma_p^{-1} Z_p}{g_0^\intercal \Sigma_p^{-1} g_0}, \quad p \in \mathcal{P}_{\mathfrak{a}_B}. \tag{16}$$

As in the literature on conditional inference (Moreira (2003), Andrews (2016a)), conditioning on the sufficient statistics for the nuisance parameters $(\xi_p)_p$ yields a condi-

tional distribution free of these nuisance parameters, thereby ensuring identification-robust inference. The second is an ancillary statistic for β under the null, which is the AR statistic corresponding to $S_{n,p}(\beta_0)$ in (13):

$$S_p \equiv h_0^\intercal Z_p = Y_p - \beta_0 X_p, \quad p \in \mathcal{P}_{\mathfrak{a}_B}.$$

The properties of the two statistics are displayed in the following lemma.

Lemma 1. In the Gaussian limit experiment (15), for each $p \in \mathcal{P}_{\mathfrak{a}_B}$, the following results hold:

(a) Under $H_0 : \delta = 0$, the profile maximum likelihood estimator T_p of ξ_p in the limit problem, defined in (16), is minimal sufficient for ξ_p and complete.

(b) Under both the null and the alternative, the two statistics S_p and T_p are independent. Specifically, their joint distribution is as follows.

$$\begin{pmatrix} S_p \\ T_p \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \delta\xi_p \\ (1+\delta\theta_p)\xi_p \end{pmatrix}, \begin{pmatrix} v_{S,p} & 0 \\ 0 & v_{T,p} \end{pmatrix} \right), \tag{17}$$

where $v_{S,p} \equiv h_0^\intercal \Sigma_p h_0$, $v_{T,p} \equiv (g_0^\intercal \Sigma_p^{-1} g_0)^{-1}$, and $\theta_p \equiv v_{T,p} g_0^\intercal \Sigma_p^{-1} e_2$.

Lemma 1(a) implies that nuisances $(\xi_p)_{p\in\mathcal{P}_{\mathfrak{a}_B}}$ cannot affect the null distribution of any tests conditional on $(T_p)_{p\in\mathcal{P}_{\mathfrak{a}_B}}$. The independence between S_p and T_p in Lemma 1(b) permits simulation of conditional critical values. Let $S_p^* \equiv \frac{S_p}{\sqrt{v_{S,p}}}$, $T_p^* \equiv \frac{T_p}{\sqrt{v_{T,p}}}$, and $\xi_p^* \equiv \frac{\xi_p}{\sqrt{v_{T,p}}}$. The standardized versions of the two statistics satisfy:

$$\begin{pmatrix} S_p^* \\ T_p^* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \delta\lambda_p\xi_p^* \\ (1+\delta\theta_p)\xi_p^* \end{pmatrix}, I_2 \right), \quad \text{independent across } p \in \mathcal{P}_{\mathfrak{a}_B} \tag{18}$$

where $\lambda_p \equiv \sqrt{v_{T,p}/v_{S,p}}$. To stack S_p^* and T_p^* across p, define $\xi^* \equiv (\xi_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}}$ and $S^* \equiv (S_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}} \sim \mathcal{N}(\delta\Lambda\xi^*, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$, $T^* \equiv (T_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}} \sim \mathcal{N}((I_{|\mathcal{P}_{\mathfrak{a}_B}|} + \delta\Theta)\xi^*, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$ where $\Lambda \equiv \mathrm{diag}((\lambda_p)_{p\in\mathcal{P}_{\mathfrak{a}_B}})$, $\theta \equiv (\theta_p)_{p\in\mathcal{P}_{\mathfrak{a}_B}}$, and $\Theta \equiv \mathrm{diag}(\theta)$. Clearly, T^* remains a sufficient statistic for $(\xi_p)_{p\in\mathcal{P}_{\mathfrak{a}_B}}$. Following Andrews, Moreira, and Stock (2006), define

$$Q \equiv (S^*\ T^*)^\intercal (S^*\ T^*) = \begin{pmatrix} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} (S_p^*)^2 & \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} S_p^* T_p^* \\ \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} S_p^* T_p^* & \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} (T_p^*)^2 \end{pmatrix} \equiv \begin{pmatrix} Q_S & Q_{ST} \\ Q_{ST} & Q_T \end{pmatrix}$$

I first derive the Lagrange multiplier (LM) statistic and its null distribution in the limit problem.

Theorem 3 (Lagrange multiplier test)**.** In the Gaussian limit experiment (15), the Lagrange multiplier statistic for $H_0 : \beta = \beta_0$ is

$$LM = \frac{\left(\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \lambda_p S_p^* T_p^*\right)^2}{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \lambda_p^2 (T_p^*)^2}. \tag{19}$$

Under H_0, $LM \mid T^* \sim \chi^2_1$, and therefore unconditionally $LM \sim \chi^2_1$.

The LM test statistic is in the same spirit as the K-statistic (Kleibergen (2002)) and score-type test (Moreira (2009)), noting that the parameter of interest β is a scalar. Let $\varphi^\alpha_{LM} \equiv 1\{LM > \chi^2_{1,1-\alpha}\}$ denote the LM test. Under strong identification, it can be shown that φ^α_{LM} is optimal in the sense of being the uniformly most powerful unbiased (UMPU) test, see Proposition S2(a) in the Supplement. Nevertheless, under weak identification, its power can be low and non-monotonic in $|\delta|$ as shown by simulation, see Figure 4. These properties of the LM test are in line with the weak IV literature, see Andrews, Moreira, and Stock (2006) and Andrews (2016a).

I next derive the conditional likelihood ratio (CLR) test in the standardized experiment generated by $(S_p^*, T_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}}$. Unlike the LM test, the CLR test remains powerful under weak identification due to the conditionality which adjusts critical values according to the signal strength reflected by the sufficient statistic T^*.

Theorem 4 (Conditional likelihood ratio test)**.** In the limit experiment (15) for testing $H_0 : \beta = \beta_0$ against $H_1 : \beta \neq \beta_0$, suppose that $|\mathcal{P}_{\mathfrak{a}_B}| \geq 2$. Then the following results hold:

(a) The likelihood ratio test statistic is

$$LR = \sup_{\delta\in\mathbb{R}} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{\left\{\delta\lambda_p S_p^* + (1+\delta\theta_p) T_p^*\right\}^2}{\delta^2\lambda_p^2 + (1+\delta\theta_p)^2} - T^{*\intercal} T^*. \tag{20}$$

Let $\zeta_p \overset{iid}{\sim} \mathcal{N}(0,1)$ for $p \in \mathcal{P}_{\mathfrak{a}_B}$. The null distribution of LR conditional on T^* is

$$LR \mid T^* \overset{\mathcal{L}}{=} F_{T^*} \equiv \sup_{\delta\in\mathbb{R}} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{\left\{\delta\lambda_p \zeta_p + (1+\delta\theta_p) T_p^*\right\}^2}{\delta^2\lambda_p^2 + (1+\delta\theta_p)^2} - T^{*\intercal} T^*,$$

where F_{T^*} is a continuously distributed random variable. The conditional test $\varphi^{\alpha}_{CLR} \equiv 1\{LR \mid T^* > cv^{\alpha}_{CLR}(T^*)\}$ in which $cv^{\alpha}_{CLR}(T^*)$ denotes the $(1-\alpha)$-quantile of F_{T^*} is conditionally similar in the sense that $\mathbb{E}_{H_0}[\varphi^{\alpha}_{CLR} \mid T^*] = \alpha$.

(b) Additionally, if $\Sigma_p = \Upsilon_p \bar{\Sigma}$ for a scalar $\Upsilon_p \in \mathbb{R}_{++}$ and a positive definite 2×2 matrix $\bar{\Sigma}$, then the likelihood ratio test statistic admits a closed form:

$$LR = LR^h = \frac{1}{2}\left(Q_S - Q_T + \sqrt{(Q_S + Q_T)^2 - 4(Q_S Q_T - Q_{ST}^2)}\right). \tag{21}$$

Conditional on T^* and under the null, the statistic LR^h is distributed as

$$\frac{1}{2}\left(Q_\perp + Q_\parallel - T^{*\intercal}T^* + \sqrt{\left(Q_\perp + Q_\parallel + T^{*\intercal}T^*\right)^2 - 4T^{*\intercal}T^* \cdot Q_\perp}\right), \tag{22}$$

where $Q_\parallel \sim \chi^2_1$ is independent of $Q_\perp \sim \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}$. Define the conditional test $\varphi^{\alpha}_{CLR^h} \equiv 1\{LR^h \mid T^* > cv^{\alpha}_{CLR^h}(T^*)\}$ where $cv^{\alpha}_{CLR^h}(T^*)$ is the $(1-\alpha)$-quantile of the conditional distribution in (22). The test $\varphi^{\alpha}_{CLR^h}$ maintains conditional similarity under heterogeneity in the sense that $\mathbb{E}_{H_0}[\varphi^{\alpha}_{CLR^h} \mid T^*] = \alpha$.

The forms of CLR tests given by Theorem 4 are similar to those in the weak-instrument literature, see (2.6) of Moreira and Moreira (2019) and (3) of Moreira (2003). The key difference lies in the asymptotic scheme: in weak IV, the limit experiment is univariate Gaussian due to the central limit theorem, while in jump regressions, the limit experiment is multivariate Gaussian induced by the semimartingale modelling.

Under strong identification, Proposition S2(a) in the supplementary material shows that the CLR test φ_{CLR} is asymptotically equivalent to the LM test φ_{LM} and therefore a UMPU test, in line with Theorem 6 in Andrews, Moreira, and Stock (2006). The CLR test under homogeneity, φ^h_{CLR}, although it is not the CLR test under heterogeneity, still controls size and lies in the class of Quasi-CLR (QCLR) tests motivated by Kleibergen (2005) in a weak GMM setting; see also Andrews (2016a). Accordingly, provided a function $r : \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \to \mathbb{R} \cup \{\infty\}$ and conditional on T^*, the class of the QCLR test statistics indexed by $r(\cdot)$ is

$$QCLR_r = \frac{1}{2}\left[Q_S - r(T^*) + \sqrt{(Q_S + r(T^*))^2 - 4(Q_S - LM) \cdot r(T^*)}\right],$$

whose null distribution is conditionally pivotal. The QCLR tests can control size but it is hard to find an optimal function $r^*(\cdot)$ for a power-maximizing criterion. Intuitively, there is an efficiency loss of $\varphi^{\alpha}_{CLR^h}$ compared to φ^{α}_{CLR} with strong jumps under the default heterogeneity setting, which is confirmed by simulations. Nevertheless, Corollary 1 below justifies the CLR^h test for weak jumps.

Corollary 1. Under Assumptions 1′, 2, 3, and conditional on T^*, the conditional tests φ^{α}_{CLR} and $\varphi^{\alpha}_{CLR^h}$ defined in Theorem 4 are admissible for every $\alpha \in (0,1)$. That is, in the weakly identified limit problem, neither φ^{α}_{CLR} nor $\varphi^{\alpha}_{CLR^h}$ dominates the other.

I then develop the CLC test of Andrews (2016a) which adaptively weights the LM and AR tests according to identification strength summarized by the sufficient statistic, thereby combining the former's efficiency under strong identification with the latter's robustness under weak identification. I transform T^* to $D \equiv \Lambda T^*$ and use D as conditioning statistic for $\mathbb{E}(D) = \frac{\partial}{\partial \delta}\mathbb{E}(S^*) \mid_{\delta=0}$ extracting identifying information along its steepest local gradient. Let $J \equiv Q_S - LM$. Conditional on D,

$$LM = S^{*\intercal} P_D S^* \sim \chi^2_1, \qquad J \sim \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}, \qquad LM \perp\!\!\!\perp J$$

where $P_D = D(D^\intercal D)^{-1} D^\intercal$. Let $\mathcal{D}$ denote the support of D and $\mathcal{A}$ be the collection of measurable functions $a : \mathcal{D} \to [0,1]$. For $a \in \mathcal{A}$, the CLC test is

$$\varphi^{\alpha}_{CLC,a(D)} = 1\{LM + a(D) \cdot J > c_\alpha(a(D))\}$$

where $c_\alpha(a(D))$ is the $(1-\alpha)$-quantile of $\chi^2_1 + a(D) \cdot \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}$ in which the two chi-squares are independent. Let $\mu_D \equiv \mathbb{E}(D) = \Lambda(I + \delta\Theta)\xi^*$ and $\Sigma_D \equiv \mathrm{Var}(D) = \Lambda^2$. Define the concentration parameter by

$$\mathfrak{c} = \mu_D^\intercal \Sigma_D^{-1} \mu_D = \|(I + \delta\Theta)\xi^*\|^2$$

Let $m \equiv \mathbb{E}(S^*) = \delta\Lambda\xi^*$. Define a compatible set $\mathcal{M}_D(\mu_D)$ for m:

$$\mathcal{M}_D(\mu_D) \equiv \left\{ m = \delta\Lambda\xi^* : \mu_D = \Lambda(I_{|\mathcal{P}_{\mathfrak{a}_B}|} + \delta\Theta)\xi^*, \delta \in \mathbb{R}, \xi^* \in \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \right\}. \tag{23}$$

Thus, $S^* \sim \mathcal{N}(m, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$, $m \in \mathcal{M}_D(\mu_D)$, $D \sim \mathcal{N}(\mu_D, \Lambda^2)$, and $S^* \perp\!\!\!\perp D$. Following Andrews (2016a), I consider the plug-in minimax regret unconditional (MMRU) test.

For the CLC test $\varphi^{\alpha}_{CLC,a}$, define a power envelope by adjusting a, i.e., $\beta^u(m,\mu_D) \equiv \sup_{a\in[0,1]} \mathbb{E}_{m,\mu_D}(\varphi^{\alpha}_{CLC,a})$. The weight is selected by

$$a_{\mathrm{MMRU}}(\mu_D) \equiv \max \arg\min_{a\in[0,1]} \sup_{m\in\mathcal{M}_D(\mu_D)} \left\{\beta^u(m,\mu_D) - \mathbb{E}_{m,\mu_D}\left[\varphi^{\alpha}_{CLC,a}\right]\right\}. \tag{24}$$

The plug-in MMRU test φ^{α}_{PI} is

$$\varphi^{\alpha}_{PI} \equiv \varphi^{\alpha}_{CLC,a_{\mathrm{MMRU}}(\hat{\mu}_D)} = 1\{LM + a_{\mathrm{MMRU}}(\hat{\mu}_D)\cdot J > c_\alpha(a_{\mathrm{MMRU}}(\hat{\mu}_D))\},$$

with a plausible estimate $\hat{\mu}_D$ for μ_D. Noting $D^\intercal\Lambda^{-2}D$ is a biased estimator for $\mathfrak{c}$ since $D^\intercal\Lambda^{-2}D \sim \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|}(\mathfrak{c})$ implies $\mathbb{E}(D^\intercal\Lambda^{-2}D) = \mathfrak{c} + |\mathcal{P}_{\mathfrak{a}_B}|$, I consider three options suggested by Andrews (2016a): the MLE $\hat{\mathfrak{c}}_{\mathrm{MLE}}$, the positive part bias-corrected estimator $\hat{\mathfrak{c}}_{\mathrm{PP}} \equiv \max(Q_T - |\mathcal{P}_{\mathfrak{a}_B}|, 0)$, and the smooth shrinkage estimator defined by

$$\hat{\mathfrak{c}}_{\mathrm{KRS}} \equiv Q_T - |\mathcal{P}_{\mathfrak{a}_B}| + e^{-\frac{Q_T}{2}}\left(\sum_{j=0}^{\infty}\left(-\frac{Q_T}{2}\right)^j \frac{1}{j!(|\mathcal{P}_{\mathfrak{a}_B}|+2j)}\right)^{-1}.$$

Accordingly, the plug-in MMRU test applies the estimated mean of D

$$\hat{\mu}_{D,i} \equiv D\cdot\sqrt{\hat{\mathfrak{c}}_i/Q_T}, \quad i\in\{\text{MLE, PP, KRS}\}.$$

The following theorem collects the necessary results.

Theorem 5 (Conditional linear combination test)**.** Under Assumptions 1′, 3, and $|\mathcal{P}_{\mathfrak{a}_B}| \geq 2$, the following results hold for the CLC test.

(a) For any measurable function $a : \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \to [0,1]$, the CLC test $\varphi^{\alpha}_{CLC,a(D)} = 1\{LM + a(D)\cdot J > c_\alpha(a(D))\}$ satisfies $\mathbb{E}_{H_0}\left[\varphi^{\alpha}_{CLC,a(D)} \mid D\right] = \alpha$.

(b) In the strong or semi-strong identification cases, the MMRU weight defined by (24) satisfies $a_{\mathrm{MMRU}}(\mu_D) \to 0$.

By Lemma 1(a), T^* is complete under H_0, so every similar test has Neyman structure with respect to T^*; restricting attention to conditionally similar tests such as $\varphi^{\alpha}_{CLC,a(D)}$ is therefore without loss of generality. Moreover, the MMRU weight converges to zero under strong or semi-strong identification, which implies that the CLC test reduces to the LM test in these cases and achieves efficency.

6 Simulation Evidence

Following Bollerslev and Todorov (2011), I adopt a two-factor volatility structure for the benchmark data generating process (DGP) with calibrated parameters, incorporating key features of high-frequency financial data, such as continuous beta, leverage effect, price-volatility co-jumps, and idiosyncratic Y-jumps. The process is set to span over one year, corresponding to 250 trading days with 6.5 daily trading hours. I discretize the process at 1-minute frequency to ameliorate the microstructure noise concern while maintaining the role of infill asymptotics. The announcements $\mathfrak{a}$ are pre-scheduled at 2 p.m. on the 20th, 60th, 80th, 110th, 130th, 160th, 205th, and 230th trading days in one year, mimicking the FOMC announcements. Let β denote the jump beta. The two factor processes (V_1, V_2) evolve as follows:

$$\begin{aligned}
\mathrm{d}V_{1,t} &= 0.0128(0.4068 - V_{1,t})\mathrm{d}t + 0.0954\sqrt{V_{1,t}}(\rho \mathrm{d}W_{1,t} + \sqrt{1-\rho^2}\mathrm{d}B_{1,t}),\\
\mathrm{d}V_{2,t} &= 0.6930(0.4068 - V_{2,t})\mathrm{d}t + 0.7023\sqrt{V_{2,t}}(\rho \mathrm{d}W_{1,t} + \sqrt{1-\rho^2}\mathrm{d}B_{2,t}),
\end{aligned}$$

where W_1, B_1, and B_2 are independent standard Brownian motions, with W_1 capturing the systematic shock and (B_1, B_2) driving idiosyncratic shocks, and $\rho = -0.7$ reflects the leverage effect, and the initial value is set to $(V_{1,0}, V_{2,0}) = (0.5, 0.5)$. The volatility processes σ_X and σ_Y depend on factors according to

$$\begin{aligned}
\sigma^2_{X,t} &= V_{1,t} + V_{2,t} + \sum_{p=1}^{8} J^{\sigma}_p \cdot 1\{\mathfrak{a}_p \leq t\}, \quad J^{\sigma}_p \overset{\text{iid}}{\sim} \exp(0.125)\\
\sigma^2_{Y,t} &= [2 - 0.75\sin^2(t)](V_{1,t} + V_{2,t})
\end{aligned}$$

where $J^{\sigma}_p, p = 1, \cdots, 8$ are i.i.d. exponential random variables with mean equal to eight. I set $Z_0 = 0$, and the dynamics of log-prices are described as follows.

$$\mathrm{d}X_t = \sigma_{X,t}\mathrm{d}W_{1,t} + \sum_{p=1}^{6} \xi_p 1\{\mathfrak{a}_p \leq t\} \cdot \Delta^r_n, \quad \xi_p \overset{\text{iid}}{\sim} \eta_p \cdot \mathcal{N}(5, 0.25),$$

$$\mathrm{d}Y_t = \beta_c \cdot \sigma_{X,t}\mathrm{d}W_{1,t} + \sigma_{Y,t}\mathrm{d}W_{2,t} + \varphi_{Y,t}\mathrm{d}N_t \cdot 1\{t \notin \mathfrak{a}\} + \beta \sum_{p=1}^{6} \xi_p 1\{\mathfrak{a}_p \leq t\} \cdot \Delta^r_n,$$

$\beta_c = 0.89, \varphi_{Y,t}|\sigma_{Y,t} \sim \mathcal{N}(0, 25 \cdot \sigma^2_{Y,t})$, N_t is a Poisson process with rate $\lambda_N = 15$.

Table 1: Monte Carlo null rejection percentages at the 5% nominal level.

r	AR	AR	LM	CLR	CLR^h	CLC-$\hat{\mathfrak{c}}_{\text{MLE}}$	CLC-$\hat{\mathfrak{c}}_{\text{PP}}$	CLC-$\hat{\mathfrak{c}}_{\text{KRS}}$	LTT–GLS	LTT–OLS
0	6.88	6.37	6.00	5.86	5.42	6.02	6.02	6.02	5.78	5.15
0.1	6.88	6.37	5.95	5.99	5.42	5.92	5.92	5.92	6.47	5.85
0.3	6.89	6.36	6.02	5.89	5.46	6.14	6.14	6.14	11.87	10.52
0.5	6.90	6.37	5.68	6.55	6.12	6.54	6.56	6.51	30.98	25.98
κ	Obs.	Unobs.	Obs.	Obs.	Obs.	Obs.	Obs.	Obs.	Unobs.	Unobs.

Notes: Entries are null rejection percentages from 10,000 replications with $\beta_0 = \beta = 1$ and $\kappa_{n,p} = 0$ in the DGP; r indexes jump shrinkage. In the κ row, Obs./Unobs. means observed/unobserved for inference. Procedures with unobserved κ use simulated critical values based on $\kappa_p \overset{\text{iid}}{\sim} \text{Uniform}(0,1)$.

where β_c is the continuous beta, and the η_p are independent Rademacher random variables indicating the jump direction. To demonstrate the robustness of the proposed methods for "no jumps", I let no jump occur at the last two announcements.

On tuning parameters, I use a balanced window for large-k volatility estimation, specifically, $k_n = 3 \times \lfloor \Delta_n^{-1/2} \rfloor = 57$. As for the truncation level $u_n = (u_n^X, u_n^Y)^\intercal$, it is necessary to scale the threshold by the surrounding volatility level, and I set

$$u_n^X = 4 \times \sqrt{\text{BV}_d} \times \Delta_n^{0.49}, \quad \text{BV}_d = \frac{\pi}{2} \frac{\lfloor 1/\Delta_n \rfloor}{\lfloor 1/\Delta_n \rfloor - 1} \sum_{i=\lfloor (d-1)/\Delta_n \rfloor + 2}^{\lfloor d/\Delta_n \rfloor} |\Delta_{i-1}^n X| |\Delta_i^n X|, \quad (25)$$

and u_n^Y is likewise calculated. Here BV_d denotes the bipower variation estimator on day d that is robust to jumps (Barndorff-Nielsen and Shephard (2004)).

Table 1 summarizes the size control results. I set $\beta = 1$ in the benchmark DGP and vary the jump shrinkage rate over $r \in \{0, 0.1, 0.3, 0.5\}$, which span strong, semi-strong, and weak identification. It is found that rejection probabilities of LTT's OLS and GLS become severely oversized at $r = 0.3$ and 0.5. Moreover, GLS is more size-distorted than OLS, reflecting the robustness and efficiency trade-off. At the same time, AR and the conditional procedures remain much closer to the nominal level as jumps weaken, although some modest upward size distortion can remain due to finite-sample error in large-k estimation; the oracle comparison in Table S4 of the Supplement supports this interpretation.

Figure 4 reports power from variants of the benchmark DGP. Three features stand out. First, the finite-sample rejection curves are asymmetric because changing β_0 changes the covariance projection used to standardize the statistics. On the positive

side of panels (a) and (b), CLR has the highest power and CLC tests remain powerful due to the numerical design, while LM's power rises much more slowly, showing its inefficiency. Second, panel (c) reproduces the power-ditch phenomenon in the weak-IV literature without changing the benchmark of price-jump signals but only adjusting the continuous beta. The nonmonotonicity is a score-cancellation effect: as β_0 moves, announcement-specific contributions to the LM score change sign and increasingly offset one another. Third, under strong identification in panel (d), LM, heterogeneous CLR, and CLC approach one another, while CLR^h exhibits the efficiency loss, as predicted by Proposition S2 in the Supplement. The AR test is inefficient in this strongly identified and overidentified setting, as expected. Overall, the proposed tests are robust to weak identification and the conditional tests maintain maximized power under strong identification. I therefore recommend the CLR test as the default choice, and the AR test in the just-identified case for detecting time variation in the parameter of interest.

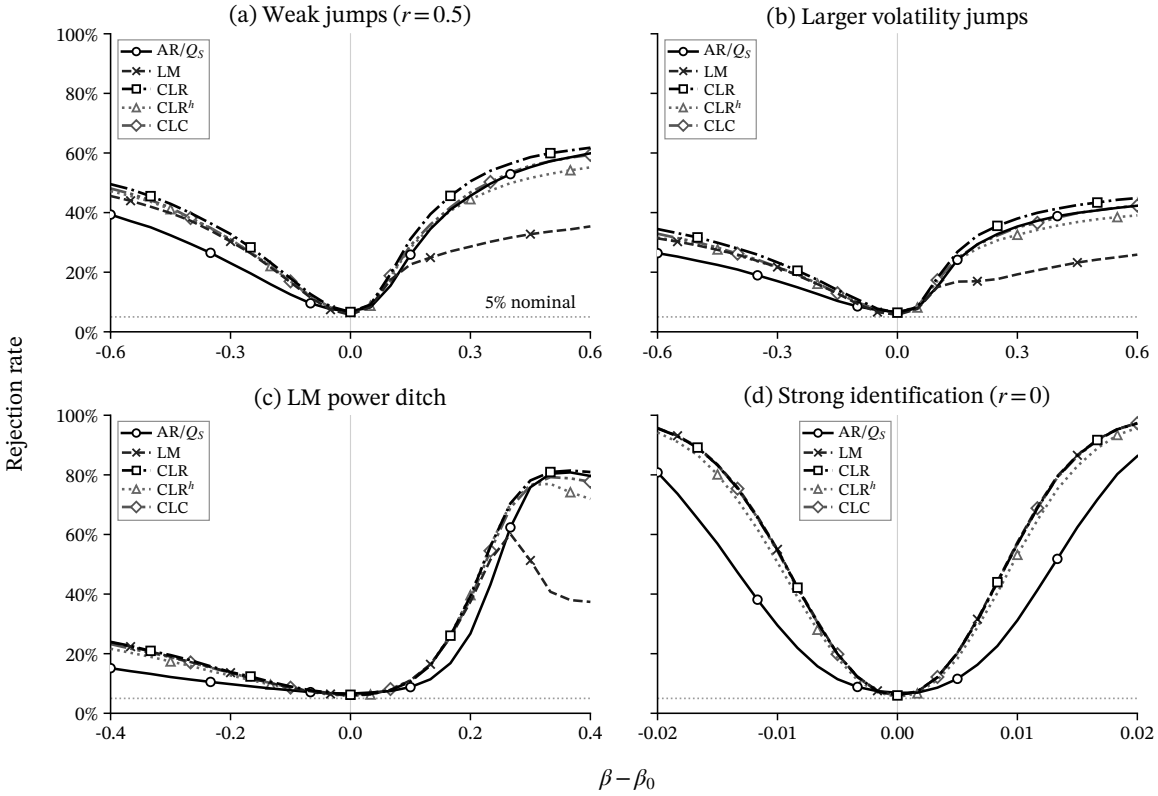
7 Empirical Applications

7.1 Infill-Asymptotic Analysis for the Fed Information Effect

The first empirical application examines announcement-specific evidence on the Fed information effect. Nakamura and Steinsson (2018) document that policy news raising interest rates is associated with upward revisions in expected output growth. They interpret this finding as evidence that FOMC announcements convey information about economic fundamentals in addition to the future path of monetary policy. Bauer and Swanson (2023a) offer an alternative explanation: monthly forecast revisions and high-frequency monetary policy surprises may both reflect previously released economic news, with the Fed responding more strongly than market participants anticipated. Golez and Matthies (2025) use dividend strip data to find evidence of the Fed information effect. Those debates reflect the difficulty of attributing lower-frequency changes in expectations to a particular announcement and the need for analysis at higher frequencies.

I use one-minute data around nine FOMC announcements from September 17, 2025 to September 16, 2026. Following the construction of Acosta, Brennan, and Jacobson (2024), I construct the policy-news factor X as the first principal component

Figure 4: Power under one-minute benchmark-DGP variants.

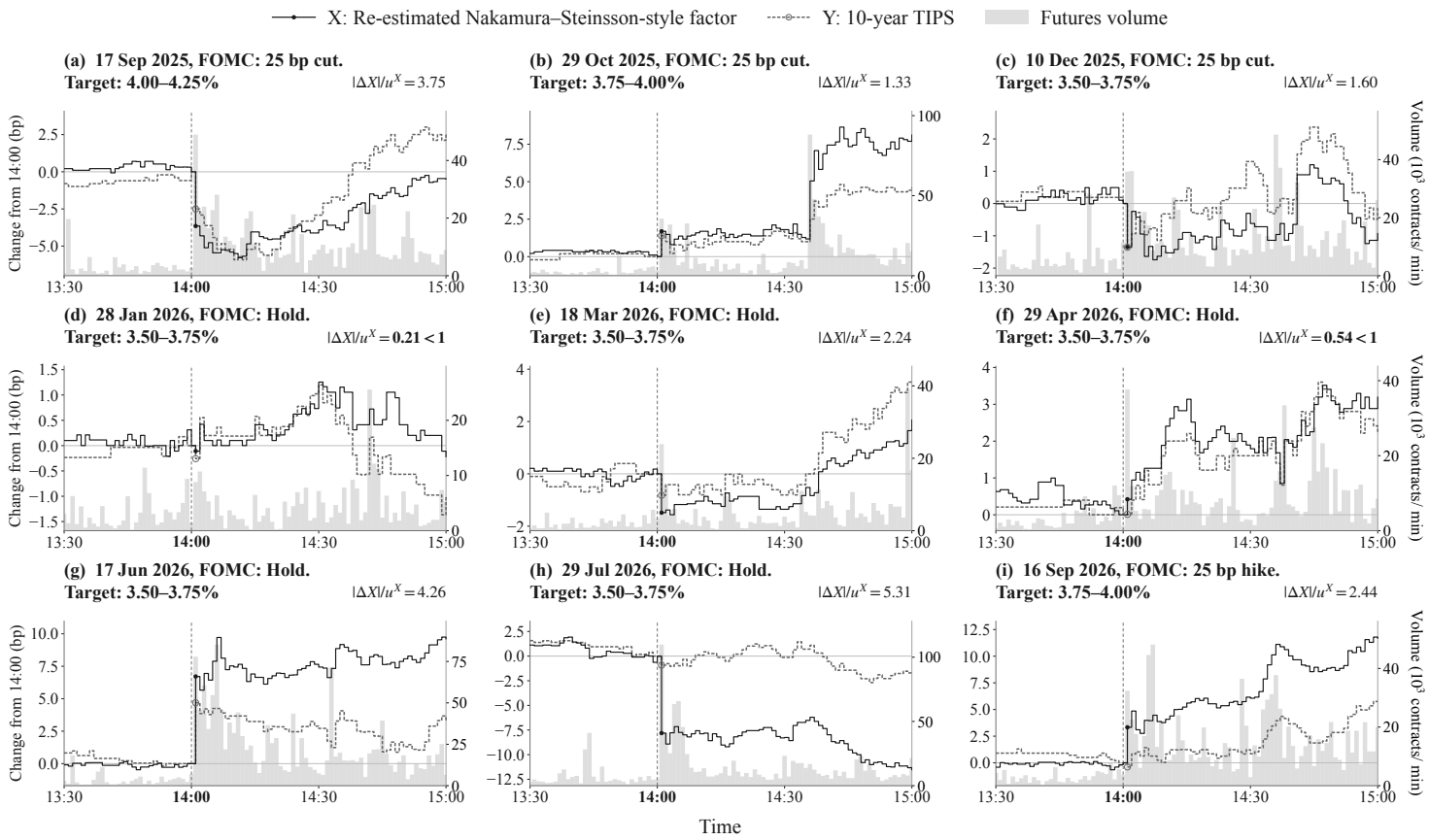


Notes: Rejection percentages use 10,000 replications from variants of the benchmark DGP with $|\mathcal{P}_{\mathfrak{a}_B}| = 8$, $\beta = 1$, $\kappa_{n,p} = 0$, and $k_n = 57$. Panels (a)–(d) use $(r, \beta_c) = (0.5, 0.89)$, $(0.5, 0.89)$, $(0.5, 0.70)$, and $(0, 0.89)$ and volatility-jump means 8, 16, 16, and 8, respectively. The dotted line marks the 5% nominal level. Across all tests, null rejection probabilities range from 5.5% to 6.9%.

of changes in the rates implied by two federal funds (FF) futures and three secured overnight financing rate (SOFR) futures, oriented so that positive changes indicate tightening news. I consider three outcome variables: Y can be (i) ten-year treasury inflation-protected securities (TIPS) midpoint yield obtained from LSEG, (ii) $100 \log D_t$, where D_t is the price of 180-day dividend claims inferred from put–call parity; or (iii) $100 \log S_t$, where S_t is the S&P 500 index. Infill asymptotics accommodates this fixed span and small number of announcements. Figure 5 diplays the one-minute intraday data of the Nakamura and Steinsson (2018)-style news policy factor and TIPS.

Table 2 displays the empirical results for 95% confidence sets obtained using identification-robust inferential methods (the 99% confidence sets can be found in Table S1), and shows substantial variation across announcements. Panel A reports

Figure 5: Policy news and real yields around nine FOMC announcements, 2025–2026



Notes: All times are New York time. X uses the first principal component of standardized 13:50–14:20 changes in rates implied by two federal funds futures and three quarterly SOFR futures across the nine announcements. I rescale its weights such that X's effect on the longest-maturity SOFR contract used here (SOFR5) is one, adapting the normalization of Nakamura and Steinsson (2018, p. 1294). The scaled weights are applied to one-minute data. Header ratios compare the 14:00–14:01 change with the threshold in equation (25). Missing prices are carried forward. Volume sums contract counts across the five futures.

Table 2: Identification-robust 95% confidence sets for one-minute responses to policy news

FOMC day	A. Real yield	B. Dividend price	C. Stock return	Method
2025-09-17	$[0.494, 0.875]$	$[-0.458, 11.086]$	$[-0.066, 0.005]$	AR
2025-10-29	$[0.539, 1.118]$	$[-13.249, -1.438]$	$[0.008, 0.078]$	AR
2025-12-10	$[0.515, 1.906]$	$[-5.111, 11.342]$	$[-0.370, -0.124]$	AR
2026-01-28	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	AR
2026-03-18	$[0.230, 0.863]$	$[-5.638, 2.832]$	$[-0.099, 0.002]$	AR
2026-04-29	$\mathbb{R} \setminus (1.303, 2.091)$	$\mathbb{R} \setminus (-5.002, 36.148)$	$\mathbb{R} \setminus (-0.061, 0.278)$	AR
2026-06-17	$[0.608, 0.824]$	$[-0.239, 4.361]$	$[-0.079, -0.039]$	AR
2026-07-29	$[0.033, 0.200]$	$[1.577, 4.811]$	$[-0.047, -0.030]$	AR
2026-09-16	$[-0.473, 0.113]$	$[-1.117, 4.323]$	$[-0.045, -0.001]$	AR
Pooled	$[0.450, 0.554]$	$[0.92, 3.00]$	$[-0.047, -0.031]$	CLR

Notes: Announcement-specific sets use AR with χ^2_1 critical values; pooled sets use CLR with a common β across 9 meetings. All panels use X with the normalization described in Figure 5. Units are bp (panel A) or log-return percentage points (panels B and C) per SOFR5-equivalent bp. Panel B uses 180-day dividend prices inferred from put–call parity; Panel C uses the S&P 500 index. Covariance estimates use $\kappa = 0$ and $k_n = 30$ (panel A) or 28 (panels B and C); PCA weights are treated as fixed.

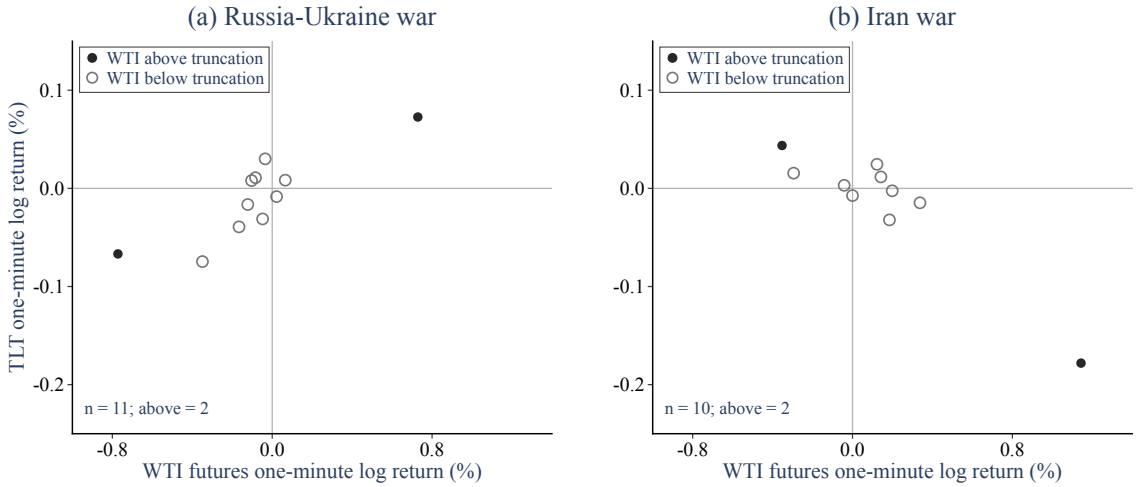
positive real-yield responses at six of the nine meetings, with confidence sets excluding zero at the 95% level, consistent with the monetary non-neutrality findings of Nakamura and Steinsson (2018). Dividend-price responses in Panel B provide complementary evidence of structural changes in the information channel (Golez and Matthies, 2025): at the 95% level, the July 2026 response is positive and the October 2025 response is negative. Panel C shows negative stock-return responses in December 2025 and June, July, and September 2026, but a positive response in October 2025, at the same confidence level. These patterns are consistent with varying contributions of information and conventional policy effects. Identification strength also varies: at the 95% level, the January sets equal $\mathbb{R}$, indicating no jumps, and the April sets are disjoint and unbounded, a case theoretically analyzed by (Mikusheva, 2010). Using AR test for the just-identified case captures the time-varying causal parameter from the data, and this empirical finding accords with Amir-Ahmadi, Matthes, and Wang (2026). Imposing a common β across meetings yields sharper CLR 95% sets of $[0.450, 0.554]$, $[0.92, 3.00]$, and $[-0.047, -0.031]$ for real yields, dividend prices, and stock returns, respectively, supporting a monetary non-neutrality and Fed informa-

tion effect in the over-identification.

7.2 Analysis of Oil Shocks by Wars

This subsection examines the relationship between war-induced oil price jumps and long-term U.S. Treasury bond prices. The explanatory variable X is the West Texas Intermediate (WTI) crude oil futures price, while the outcome variable Y is the price of the iShares 20+ Year Treasury Bond ETF (TLT), which tracks long-term U.S. Treasury bonds. We focus on news events associated with the Russia–Ukraine and Iran wars that affect global market expectations about oil supply. Detailed information on these news events is provided in Section S4.2.

Figure 6: WTI and TLT log returns around war-related oil news



Notes: One-minute log returns in percent around Russia–Ukraine war, talks, and security news (panel a) and Iran war news (panel b). Filled black circles and open gray circles denote WTI returns above and below truncation, respectively. Each panel reports event counts; 95% confidence intervals are in Table 3.

Interpreting upward WTI jumps as adverse oil-supply news, Figure 6 suggests two distinct bond-market responses. For Russia–Ukraine events, TLT prices rise with oil prices, consistent with a flight to safety as geopolitical risk increases demand for U.S. Treasuries. For Iran-war events, the negative relation is consistent with an inflation channel: expected supply losses from the blockade of the Strait of Hormuz raise inflation expectations and long-term nominal yields, depressing TLT prices. The opposite signs suggest that the balance between safe-haven demand and inflation concerns differs across the two groups of events.

Table 3 shows that each least-squares procedure yields a 95% interval containing zero in at least one sample: LTT–OLS for Russia–Ukraine, and OLS–HC3 with or without an intercept for Iran. CLR excludes zero in both, delivering a positive jump beta for Russia–Ukraine and a negative one for Iran. Identification-robust inference thus yields sharp conclusions about the direction of bond-price responses, consistent with the contrasting economic mechanisms above.

Table 3: 95% confidence intervals for war-related oil-news jump betas

News group	OLS–HC3		LTT–OLS	CLR
	With intercept	Without intercept		
Russia–Ukraine	$[0.067, 0.141]$	$[0.067, 0.143]$	$[-0.046, 0.256]$	$[0.011, 0.401]$
Iran	$[-0.279, 0.004]$	$[-0.272, 0.007]$	$[-0.163, -0.101]$	$[-0.181, -0.123]$

Notes: The OLS-HC3 uses heteroskedasticity-consistent standard errors. LTT–OLS is the jump regression and CLR is the robust test of this paper, both using observed-κ variance estimation.

7.3 Jump Beta between ETFs with FOMC Announcements

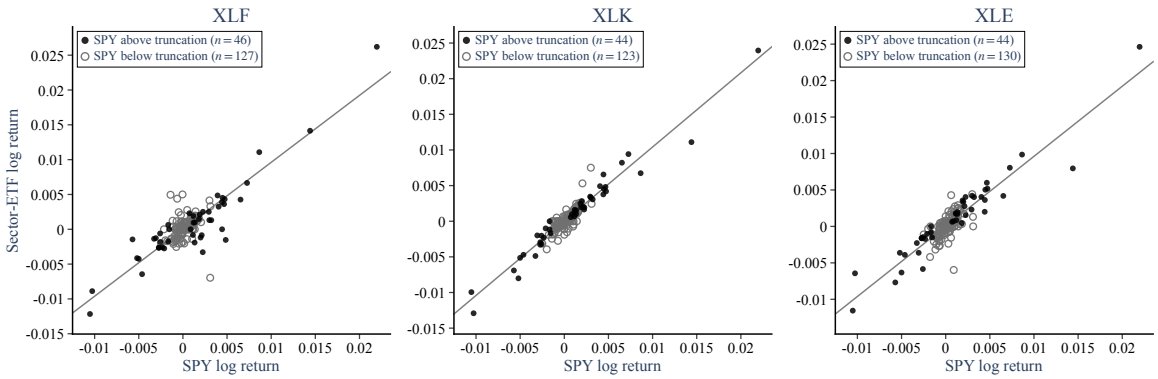
The third empirical application studies jump dependence across exchange-traded funds (ETFs) around FOMC announcements through a factor-model lens. The sample spans 2001–2024 and includes 190 scheduled FOMC meetings. While some announcements generate strongly identified jumps, a substantial fraction are characterized by small jumps.

I specify the X-process as the ETF tracking the S&P 500 index (ticker: SPY), serving as the proxy for market-wide jumps. For the Y-process, I draw from industry-specific ETFs in order to study their exposures to market jump risk. The Y-variable candidates, with corresponding tickers in parentheses, are: materials (XLB), financials (XLF), technology (XLK), and energy (XLE). I obtain the one-minute intraday data from Pi Trading[5], so that microstructure frictions such as noise and rounding errors are less of a concern. The exact announcement times of FOMC releases are available on Bloomberg[6], which allows me to precisely anchor the timing of news-jumps. Figure 7 presents scatter plots of the high-frequency return pairs around

[5]Pi Trading website: `https://pitrading.com/`.

[6]The release times of FOMC announcements can be found in Bloomberg at "FOMC Rate Decision (Upper Bound)" in the Monetary Sector, Economic Releases, Economic Calendars (United States).

Figure 7: SPY and sector-ETF log returns around FOMC announcements

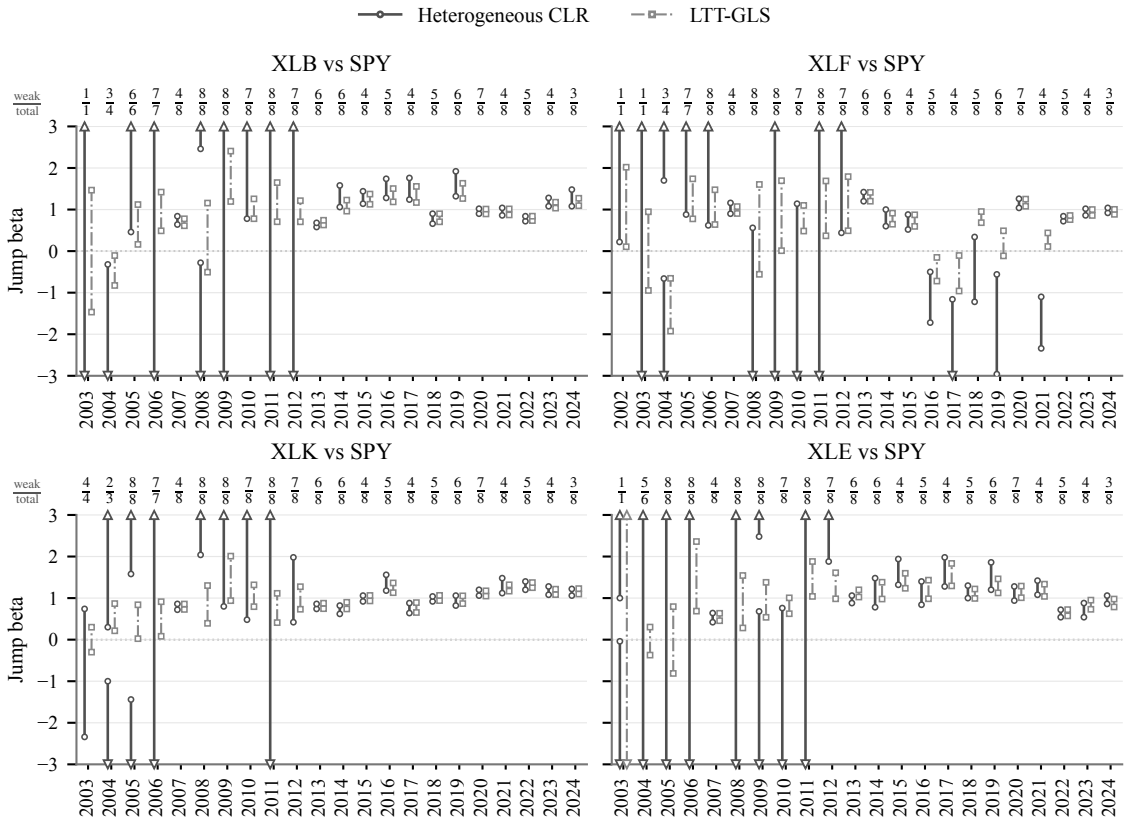


Notes: Straight lines are zero-intercept fits to the observations whose SPY log returns are above truncation.

FOMC announcements for XLF, XLK, and XLE; the corresponding plot for XLB is shown in Figure 2.

Figure 8 compares annual 95% confidence sets obtained from CLR and the LTT–GLS regression, imposing a common jump beta within each calendar year. Both methods are applied to the same set of retained announcements and use the large-k spot covariance estimator with $k_n = 20$. When weak jumps dominate, the CLR confidence sets are often wide, whereas LTT–GLS can still produce relatively tight intervals. The two methods nevertheless exhibit substantial agreement in 2007 across all sector ETFs, when 4 of the 8 announcements are strongly identified, as well as during 2013–2024 for XLB, XLK, and XLE. In the remaining cases, however, CLR often yields substantially wider confidence sets, which may even fail to overlap with those produced by LTT–GLS. The discrepancy is particularly pronounced over 2008–2012, a period in which all SPY jumps can be weakly identified. These findings illustrate that conventional least-squares-based inference can appear spuriously precise under weak identification and thereby lead to potentially misleading empirical conclusions. I therefore recommend identification-robust inference for empirical applications in which announcement-induced jumps may be weak.

Figure 8: Annual FOMC jump-beta confidence sets: heterogeneous CLR and LTT–GLS



Notes: Annual 95% confidence sets for the jump beta around FOMC announcements. Dark-gray solid lines with open circles denote CLR; light-gray dash-dotted lines with open squares denote LTT–GLS. Spot covariance estimates use $k_n = 20$. The CLR test, see (20), uses known-κ covariance estimates $\kappa_{n,p}\hat{c}_{n,i(p)-} + (1-\kappa_{n,p})\hat{c}_{n,i(p)+}$ and a 0.02 grid over $[-3, 3]$. For both methods, outward open triangles mark truncation at the display limits, not necessarily unbounded confidence sets. Fractions at the top of each panel report weak/total announcements in the corresponding year.

8 Conclusion

High-frequency event-study regressions have become increasingly popular as high-frequency identification has emerged as an important tool for causal inference in empirical macroeconomics. However, the 30-minute or longer event windows commonly used in practice may still be too wide to credibly maintain exogeneity, thereby weakening the causal interpretation of the resulting estimates (Gürkaynak, Swanson, and Wright (2026)). This paper proposes an alternative framework that exploits minute-level data through infill asymptotics and focuses on jumps, a natural object in continuous-time models, as the source of identifying variation.

A key challenge is that jumps can be small when the announcement information is largely anticipated by financial markets, giving rise to the weak identification problem. I develop inference robust to such low signals by drawing on the weak-instrument literature, while also characterizing efficient tests under strong identification. The infill-asymptotic framework is attractive in three respects: it minimizes omitted-variable bias concerns, allows informative inference even when the number of events is small, and accommodates time variation in the parameter of interest.

Appendices

Appendix A. Testing Procedures for Observed κ

This appendix details the testing procedures for the AR, LM, CLR, CLR^h, and CLC tests, given the sample $\{\Delta_i^n Z = (\Delta_i^n X, \Delta_i^n Y)^\intercal\}_i$ when Assumption 1′ is satisfied. Fix a candidate value β_0 and write $g_0 = (1, \beta_0)^\intercal$, $h_0 = (-\beta_0, 1)^\intercal$, and $e_2 = (0, 1)^\intercal$. The procedures below presume $|\mathcal{P}_{\mathfrak{a}_B}| \geq 2$; when $|\mathcal{P}_{\mathfrak{a}_B}| = 1$, use the AR test directly. At each announcement $p \in \mathcal{P}_{\mathfrak{a}_B}$, compute the spot covariance estimate $\hat{\Sigma}_p$,

$$\hat{\Sigma}_p = \kappa_{n,p}\hat{c}_{n,i(p)-} + (1 - \kappa_{n,p})\hat{c}_{n,i(p)+}, \qquad \text{where } \kappa_{n,p} = \frac{\mathfrak{a}_p - (i(p) - 1)\Delta_n}{\Delta_n}.$$

Suppressing dependence on β_0 in the following display, define

$$\begin{cases} \hat{v}_{S,p} \equiv h_0^\intercal \hat{\Sigma}_p h_0, & \hat{v}_{T,p} \equiv \left(g_0^\intercal \hat{\Sigma}_p^{-1} g_0\right)^{-1}, \\ S_{n,p} \equiv \Delta_n^{-1/2} h_0^\intercal \Delta_{i(p)}^n Z, & T_{n,p} \equiv \Delta_n^{-1/2} \hat{v}_{T,p} g_0^\intercal \hat{\Sigma}_p^{-1} \Delta_{i(p)}^n Z, \\ \hat{\lambda}_p \equiv \sqrt{\hat{v}_{T,p}/\hat{v}_{S,p}}, & \hat{\theta}_p \equiv \hat{v}_{T,p} g_0^\intercal \hat{\Sigma}_p^{-1} e_2, \\ S_{n,p}^* \equiv \hat{v}_{S,p}^{-1/2} S_{n,p}, & T_{n,p}^* \equiv \hat{v}_{T,p}^{-1/2} T_{n,p}. \end{cases} \tag{26}$$

Stack these coordinates across announcements and define

$$\begin{cases} S_n^*(\beta_0) \equiv (S_{n,p}^*(\beta_0))_p, & T_n^*(\beta_0) \equiv (T_{n,p}^*(\beta_0))_p, \\ \hat{\Lambda}_n(\beta_0) \equiv \mathrm{diag}((\hat{\lambda}_p(\beta_0))_p), & \hat{\Theta}_n(\beta_0) \equiv \mathrm{diag}((\hat{\theta}_p(\beta_0))_p), \\ D_n(\beta_0) \equiv \hat{\Lambda}_n(\beta_0) T_n^*(\beta_0), & Q_{S,n}(\beta_0) \equiv S_n^*(\beta_0)^\intercal S_n^*(\beta_0), \\ Q_{T,n}(\beta_0) \equiv T_n^*(\beta_0)^\intercal T_n^*(\beta_0), & Q_{ST,n}(\beta_0) \equiv S_n^*(\beta_0)^\intercal T_n^*(\beta_0). \end{cases} \tag{27}$$

For compactness, the dependence of the quantities on β_0 is suppressed below. For $s, t \in \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|}$, define the following likelihood-ratio functional

$$\mathcal{L}(s,t) \equiv \max_{\varrho \in [-\pi/2, \pi/2]} \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{\left\{\hat{\lambda}_p s_p \sin\varrho + (\cos\varrho + \hat{\theta}_p \sin\varrho) t_p\right\}^2}{\hat{\lambda}_p^2 \sin^2\varrho + (\cos\varrho + \hat{\theta}_p \sin\varrho)^2} - t^\intercal t, \tag{28}$$

where $\delta = \tan\varrho$ gives the reparametrization of the likelihood ratio in Theorem 4, and the reparametrization enables the feasibility of computation. Also define

$$\mathcal{L}^h(q_S, q_T, q_{ST}) \equiv \frac{1}{2}\left\{q_S - q_T + \sqrt{(q_S + q_T)^2 - 4(q_S q_T - q_{ST}^2)}\right\}. \tag{29}$$

Algorithm 2 (Consolidated testing procedures)**.** For a $\beta_0 \in \mathbb{R}$ and $\alpha \in (0,1)$:

(i) Compute the common quantities in (26)–(27) once. Set

$$LM_n \equiv \frac{(S_n^{*\intercal} D_n)^2}{D_n^\intercal D_n}, \qquad\qquad J_n \equiv Q_{S,n} - LM_n,$$

with the computational convention $LM_n = 0$ when $D_n = 0$.

(ii) Compute the statistic and its corresponding critical value for the selected test:

a. (AR) The AR test uses $Q_{S,n}$ as the test statistic, with critical value $\chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|,1-\alpha}$.

b. (LM) The LM test uses LM_n as the test statistic and the critical value $\chi^2_{1,1-\alpha}$.

c. (CLR) Calculate the test statistic:

$$LR_n \equiv \mathcal{L}(S_n^*, T_n^*),$$

where $\mathcal{L}$ is defined in (28). Independently of the data, draw $\tilde{\zeta}^{(m)} \sim \mathcal{N}(0, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$, $m = 1, \ldots, M$, and set

$$\widetilde{LR}_n^{(m)} \equiv \mathcal{L}(\tilde{\zeta}^{(m)}, T_n^*).$$

The conditional critical value $cv_{n,CLR}^{\alpha}(T_n^*)$ is the empirical $(1-\alpha)$-quantile of $(\widetilde{LR}_n^{(m)})_{m=1}^M$.

d. (CLR^h) Calculate the test statistic:

$$LR_n^h \equiv \mathcal{L}^h(Q_{S,n}, Q_{T,n}, Q_{ST,n}),$$

where $\mathcal{L}^h$ is defined in (29). Independently of the data, draw $\tilde{\zeta}^{(m)} \sim \mathcal{N}(0, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$, $m = 1, \ldots, M$. For each m, calculate

$$\tilde{Q}_{S,n}^{(m)} \equiv \tilde{\zeta}^{(m)\intercal}\tilde{\zeta}^{(m)}, \quad \tilde{Q}_{ST,n}^{(m)} \equiv \tilde{\zeta}^{(m)\intercal}T_n^*, \quad \widetilde{LR}_n^{h,(m)} \equiv \mathcal{L}^h(\tilde{Q}_{S,n}^{(m)}, Q_{T,n}, \tilde{Q}_{ST,n}^{(m)}).$$

Let $cv_{n,CLR^h}^{\alpha}(T_n^*)$ be the empirical $(1-\alpha)$-quantile of $(\widetilde{LR}_n^{h,(m)})_{m=1}^M$.

e. (CLC) For a plug-in rule $i \in \{\text{MLE}, \text{PP}, \text{KRS}\}$, form

$$\hat{\mu}_{D,n,i} \equiv \begin{cases} D_n\sqrt{\hat{\mathfrak{c}}_i/Q_{T,n}}, & Q_{T,n} > 0, \\ 0, & Q_{T,n} = 0. \end{cases}$$

where, letting $f_{\chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|}(\mathfrak{c})}$ denote the density of a noncentral chi-square distribution with $|\mathcal{P}_{\mathfrak{a}_B}|$ degrees of freedom and noncentrality parameter $\mathfrak{c}$,

$$\hat{\mathfrak{c}}_{\text{MLE}} \in \arg\max_{\mathfrak{c}\geq 0} f_{\chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|}(\mathfrak{c})}(Q_{T,n}), \quad \hat{\mathfrak{c}}_{\text{PP}} \equiv (Q_{T,n} - |\mathcal{P}_{\mathfrak{a}_B}|)_+,$$
$$\hat{\mathfrak{c}}_{\text{KRS}} \equiv Q_{T,n} - |\mathcal{P}_{\mathfrak{a}_B}| + e^{-\frac{Q_{T,n}}{2}}\left(\sum_{j=0}^{\infty}\left(-\frac{Q_{T,n}}{2}\right)^j \frac{1}{j!(|\mathcal{P}_{\mathfrak{a}_B}| + 2j)}\right)^{-1}.$$

Replace (Λ, Θ, μ_D) in (23)–(24) with $(\hat{\Lambda}_n, \hat{\Theta}_n, \hat{\mu}_{D,n,i})$ and denote the resulting largest minimizer by $\hat{a}_{n,i} \in [0,1]$; when $Q_{T,n} = 0$, set $\hat{a}_{n,i} = 1$. The statistic is

$$CLC_{n,i} \equiv LM_n + \hat{a}_{n,i}J_n$$

with critical value being the $(1-\alpha)$-quantile of $\chi_1^2 + \hat{a}_{n,i}\chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}$.

(iii) Reject $H_0 : \beta = \beta_0$ when the statistic exceeds its corresponding critical value.

Appendix B. Proofs of Main Results

Throughout, let ϕ_v denote the density function of $\mathcal{N}(0, v)$ and drop the subscript when $v = 1$.

Proof of Theorem 1

(a) Let X^c and J_X respectively denote the continuous part and the jump part of the X-process. Set $|\mathcal{I}_n(\mathfrak{a}_B)| = q < \infty$ where $\mathcal{I}_n(\mathfrak{a}_B)$ is defined in (11). Note that q is

fixed as $n \to \infty$. Then for any strictly positive integer q,

$$\begin{aligned}
\mathbb{P}(\mathcal{I}_n(\mathfrak{a}_B) \neq \varnothing) = \mathbb{P}(\exists i, i \in \mathcal{I}_n(\mathfrak{a}_B), \text{ s.t. } |\Delta_i^n X| > u_n) &= \mathbb{P}\left(\bigcup_{i \in \mathcal{I}_n(\mathfrak{a}_B)} \{|\Delta_i^n X| > u_n\} \right) \\
&\leq \sum_{i \in \mathcal{I}_n(\mathfrak{a}_B)} \mathbb{P}(|\Delta_i^n X^c + \Delta_i^n J_X| > u_n) \\
&\leq \sum_{i \in \mathcal{I}_n(\mathfrak{a}_B)} \mathbb{P}(|\Delta_i^n X^c| > u_n/2) + \sum_{i \in \mathcal{I}_n(\mathfrak{a}_B)} \mathbb{P}(|\Delta_i^n J_X| > u_n/2) \\
&\overset{(i)}{\leq} K_q \Delta_n^{q/2}/u_n^q + \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \mathbb{P}(|\xi_p| > u_n \Delta_n^{-1/2}/2) \longrightarrow 0
\end{aligned}$$

where the inequality (i) is due to Markov's inequality, $\mathbb{E}|\Delta_i^n X^c|^q \leq K_q \Delta_n^{q/2}$ for some constant K_q that depends on q, and $\mathcal{P}_{\mathfrak{a}_B}$ is a finite set.
(b) Since $R_{n,p} = \Delta_n^{-1/2}(\Delta_{i(p)}^n Z - \Delta Z_{\mathfrak{a}_p})$, $p \in \mathcal{P}_{\mathfrak{a}_B}$, it follows that

$$\begin{aligned}
&\Delta_n^{-1/2}\left(\Delta_{i(p)}^n Y - \beta \Delta_{i(p)}^n X\right) = (-\beta, 1) R_{n,p} \\
&\frac{\Delta_{i(p)}^n X}{\Delta_n^{1/2}} = \frac{\Delta_{i(p)}^n X - \Delta X_{\mathfrak{a}_p} + \Delta X_{\mathfrak{a}_p}}{\Delta_n^{1/2}} = (1,0) R_{n,p} + \xi_p
\end{aligned}$$

Therefore,

$$\begin{aligned}
\tilde{\beta}_n - \beta = \frac{\sum_{i \in \mathcal{I}(\mathfrak{a}_B)} \Delta_i^n X \Delta_i^n Y}{\sum_{i \in \mathcal{I}(\mathfrak{a}_B)} (\Delta_i^n X)^2} - \beta &= \frac{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \Delta_{i(p)}^n X (\Delta_{i(p)}^n Y - \beta \Delta_{i(p)}^n X)}{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} (\Delta_{i(p)}^n X)^2} \\
&= \frac{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} ((1,0) R_{n,p} + \xi_p)(-\beta, 1) R_{n,p}}{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} ((1,0) R_{n,p} + \xi_p)^2} \\
&\xrightarrow{\mathcal{L}\text{-}s} \frac{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} [(1,0) R_p + \xi_p] \varsigma_p(\beta)}{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} [(1,0) R_p + \xi_p]^2} = \tilde{b}
\end{aligned} \tag{30}$$

where the stable convergence in law is by (31) in the proof of Theorem 2 and the continuous mapping theorem. Finally, see Lemma 3 for a proof that $\tilde{b} = O_{\mathbb{P}}(1)$.
(c) For the weighted least squares, note that

$$\begin{aligned}
\hat{\beta}_n(\mathfrak{a}_B, w) - \beta = \frac{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} w(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}, \tilde{\beta}_n) \Delta_{i(p)}^n X \left(\Delta_{i(p)}^n Y - \beta \Delta_{i(p)}^n X\right)}{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} w(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}, \tilde{\beta}_n) \Delta_{i(p)}^n X^2} \\
\xrightarrow{\mathcal{L}\text{-}s} \frac{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} w(c_{\mathfrak{a}_p -}, c_{\mathfrak{a}_p}, \beta + \tilde{b}) [(1,0) R_p + \xi_p] \varsigma_p(\beta)}{\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} w(c_{\mathfrak{a}_p -}, c_{\mathfrak{a}_p}, \beta + \tilde{b}) [(1,0) R_p + \xi_p]^2} = b_w
\end{aligned}$$

where the stable convergence in law is due to (31), (32), (30), w being a continuous function, and the continuous mapping theorem.

Proof of Theorem 2

(a) Note $\Delta Y_{\tau_p} = \beta \Delta X_{\tau_p}$; thus $\Delta Y_{\tau_p} - \Delta^n_{i(p)} Y + \Delta^n_{i(p)} Y = \beta(\Delta X_{\tau_p} - \Delta^n_{i(p)} X + \Delta^n_{i(p)} X)$. Define $R_{n,p} \equiv \Delta_n^{-1/2}(\Delta^n_{i(p)} Z - \Delta Z_{\tau_p})$. Then $\Delta^n_{i(p)} Y = \beta \Delta^n_{i(p)} X + (-\beta, 1)\Delta_n^{1/2} R_{n,p}$. So

$$\begin{aligned}
S_{n,p}(\beta_0) &= \Delta_n^{-1/2}(\Delta^n_{i(p)} Y - \beta_0 \Delta^n_{i(p)} X) = (\beta - \beta_0)\frac{\Delta^n_{i(p)} X}{\Delta_n^{1/2}} + (-\beta, 1) R_{n,p} \\
&= (\beta - \beta_0)\frac{\Delta^n_{i(p)} X - \Delta X_{\tau_p}}{\Delta_n^{1/2}} + (\beta - \beta_0)\frac{\Delta X_{\tau_p}}{\Delta_n^{1/2}} + (-\beta, 1) R_{n,p} \\
&\overset{(i)}{=} (\beta - \beta_0)(1, 0) R_{n,p} + (\beta - \beta_0)\xi_p + (-\beta, 1) R_{n,p} \\
&= (-\beta_0, 1) R_{n,p} + (\beta - \beta_0)\xi_p
\end{aligned}$$

where equality (i) is due to Assumption 2 and the definition of $R_{n,p}$. Proposition 4.4.10 of Jacod and Protter (2012) gives the following joint convergence:

$$(R_{n,p})_{p \geq 1} \xrightarrow{\mathcal{L}\text{-}s} (R_p)_{p \geq 1} \tag{31}$$

where R_p is defined in (6). And it implies $(R_{n,p}, \xi_p) \xrightarrow{\mathcal{L}\text{-}s} (R_p, \xi_p)$ due to the property of stable convergence. According to the continuous mapping theorem,

$$S_{n,p}(\beta_0) \xrightarrow{\mathcal{L}\text{-}s} (-\beta_0, 1) R_p + (\beta - \beta_0)\xi_p = \varsigma_p(\beta_0) + (\beta - \beta_0)\xi_p$$

(b) Note that $\mathbb{E}(R_p | \sigma_{\mathfrak{a}_p -}, \sigma_{\mathfrak{a}_p}) = 0$ and $\mathrm{Var}(R_p | \sigma_{\mathfrak{a}_p -}, \sigma_{\mathfrak{a}_p}) = (c_{\mathfrak{a}_p -} + c_{\mathfrak{a}_p})/2$ by the law of iterated expectations. Hence $\mathrm{Var}(\varsigma_p(\beta_0) | c_{\mathfrak{a}_p -}, c_{\mathfrak{a}_p}) = \frac{1}{2}(-\beta_0, 1)(c_{\mathfrak{a}_p -} + c_{\mathfrak{a}_p})(-\beta_0, 1)^\intercal = v_{S,p}(\beta_0)$. By Theorem 9.3.2 in Jacod and Protter (2012), it follows that

$$\hat{c}_{n,i(p)+} \xrightarrow{\mathbb{P}} c_{\mathfrak{a}_p} \qquad \hat{c}_{n,i(p)-} \xrightarrow{\mathbb{P}} c_{\mathfrak{a}_p -} \tag{32}$$

and the continuous mapping theorem implies $\hat{v}_{S,p}(\beta_0) \xrightarrow{\mathbb{P}} v_{S,p}(\beta_0)$. By the stable convergence in law under the null in Theorem 2(a), it follows that

$$Q_S \xrightarrow{\mathcal{L}\text{-}s} \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{[\varsigma_p(\beta_0)]^2}{v_{S,p}(\beta_0)}.$$

(c) Let $cv_1^\alpha(\beta_0)$ denote the $(1-\alpha)$ quantile of $\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}[\varsigma_p(\beta_0)]^2/v_{S,p}(\beta_0)$. This quantile satisfies $cv_1^\alpha(\beta_0)>0$ for $\alpha\in(0,1)$. For the large-k estimators in (32), let $\mathbb{N}_1$ be an arbitrary subsequence of $\mathbb{N}$. Hence $(\hat{c}_{n,i(p)+},\hat{c}_{n,i(p)-})\xrightarrow{\mathbb{P}}(c_{\mathfrak{a}_p},c_{\mathfrak{a}_p-})$ along $\mathbb{N}_1$, so there is a sub-subsequence $\mathbb{N}_2$ of $\mathbb{N}_1$, s.t. $(\hat{c}_{n,i(p)+},\hat{c}_{n,i(p)-})\xrightarrow{a.s.}(c_{\mathfrak{a}_p},c_{\mathfrak{a}_p-})$ along $\mathbb{N}_2$, equivalently, the set $\tilde{\Omega}=\{\omega\in\Omega:(\hat{c}_{n,i(p)+}(\omega),\hat{c}_{n,i(p)-}(\omega))\to(c_{\mathfrak{a}_p}(\omega),c_{\mathfrak{a}_p-}(\omega))$ along $\mathbb{N}_2\}$ satisfies $\mathbb{P}(\tilde{\Omega})=1$. For fixed $\omega\in\tilde{\Omega}$, $\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}[\tilde{\varsigma}_{n,p}(\beta_0)]^2/\hat{v}_{S,p}(\beta_0)\xrightarrow{\mathcal{L}}\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}[\varsigma_p(\beta_0)]^2/v_{S,p}(\beta_0)$, as $n\to\infty$, where $\tilde{\varsigma}_{n,p}(\beta_0)$ is defined in Algorithm 1. Notice that the distribution of $\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}[\varsigma_p(\beta_0)]^2$ is atomless over $[0,\infty)$, so by Lemma 21.2 of Van der Vaart (2000) and $cv_{n,1}^\alpha(\beta_0)$ being the $(1-\alpha)$-quantile of $\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}[\tilde{\varsigma}_{n,p}(\beta_0)]^2/\hat{v}_{S,p}(\beta_0)$, it follows that $cv_{n,1}^\alpha(\beta_0,\omega)\longrightarrow cv_1^\alpha(\beta_0)$ as $n\to\infty$. Hence, for an arbitrary subsequence $\mathbb{N}_1$, there exists a subsubsequence $\mathbb{N}_2$ along which $cv_{n,1}^\alpha(\beta_0)\xrightarrow{a.s.}cv_1^\alpha(\beta_0)$. By the subsequence characterization of convergence in probability, it follows that $cv_{n,1}^\alpha(\beta_0)\xrightarrow{\mathbb{P}}cv_1^\alpha(\beta_0)$ along $\mathbb{N}$; therefore, under the null,

$$\big(Q_S,cv_{n,1}^\alpha(\beta_0)\big)\xrightarrow{\mathcal{L}\text{-s}}\Big(\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\frac{[\varsigma_p(\beta_0)]^2}{v_{S,p}(\beta_0)},cv_1^\alpha(\beta_0)\Big)$$

hence by continuous mapping, $Q_S-cv_{n,1}^\alpha(\beta_0)\xrightarrow{\mathcal{L}\text{-s}}\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}[\varsigma_p(\beta_0)]^2/v_{S,p}(\beta_0)-cv_1^\alpha(\beta_0)$, which implies $\mathbb{P}(Q_S>cv_{n,AR}^\alpha(\beta_0))\to\mathbb{P}(\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}[\varsigma_p(\beta_0)]^2/v_{S,p}(\beta_0)>cv_{AR}^\alpha(\beta_0))=\alpha$.

Proof of Lemma 1

(a) Under $H_0:\beta=\beta_0$, $Z_p\sim\mathcal{N}(g_0\xi_p,\Sigma_p)$, so the log-likelihood of Z_p in terms of ξ_p is, up to constants, $\ell(\xi_p)=-\frac{1}{2}Z_p^\intercal\Sigma_p^{-1}Z_p+\xi_p g_0^\intercal\Sigma_p^{-1}Z_p-\frac{1}{2}\xi_p^2 g_0^\intercal\Sigma_p^{-1}g_0$. Maximizing $\ell(\xi_p)$ over ξ_p generates its MLE under H_0:

$$g_0^\intercal\Sigma_p^{-1}Z_p-\hat{\xi}_p g_0^\intercal\Sigma_p^{-1}g_0=0\quad\Rightarrow\quad\hat{\xi}_p=\frac{g_0^\intercal\Sigma_p^{-1}Z_p}{g_0^\intercal\Sigma_p^{-1}g_0}=T_p$$

Furthermore, the density of Z_p can be written as

$$f_{\xi_p}(Z_p)=(2\pi)^{-1}|\Sigma_p|^{-1/2}\exp\left[-\frac{1}{2}Z_p^\intercal\Sigma_p^{-1}Z_p+\left(\xi_p\frac{g_0^\intercal\Sigma_p^{-1}Z_p}{g_0^\intercal\Sigma_p^{-1}g_0}-\frac{\xi_p^2}{2}\right)g_0^\intercal\Sigma_p^{-1}g_0\right]$$

By the factorization theorem (Corollary 2.6.1 of Lehmann and Romano (2005)), T_p is sufficient for ξ_p. And note for $Z_p \neq Z_p'$ corresponding to T_p and T_p', if

$$\log \frac{f_{\xi_p}(Z_p)}{f_{\xi_p}(Z_p')} = \xi_p g_0^\intercal \Sigma_p^{-1}(Z_p - Z_p') - \frac{1}{2}(Z_p^\intercal \Sigma_p^{-1} Z_p - Z_p'^\intercal \Sigma_p^{-1} Z_p')$$

does not depend on ξ_p, then $\xi_p g_0^\intercal \Sigma_p^{-1}(Z_p - Z_p') = 0$, implying $T_p = T_p'$. So T_p is minimal sufficient. Finally, note that $T_p \sim \mathcal{N}(\xi_p, v_{T,p})$. Suppose that $\mathbb{E}_{\xi_p}(f(T_p)) = 0, \forall \xi_p$, i.e.,

$$\int_{t\in\mathbb{R}} f(t)\phi_{v_{T,p}}(t-\xi_p)\mathrm{d}t = 0, \forall \xi_p \in \mathbb{R} \iff f * \phi_{v_{T,p}}(\xi_p) = 0.$$

Taking Fourier transforms gives $\hat{f}(\omega)\hat{\phi}_{v_{T,p}}(\omega) = 0$. Note that $\hat{\phi}_{v_{T,p}}(\omega) = e^{-v_{T,p}\omega^2/2}$; thus $\hat{f} = 0$, implying $f = 0$ a.e. over the distribution family of T_p. So T_p is complete. (b) Note $h_0 = h + \delta e_1$ where $\delta = \beta - \beta_0$, so recalling (15) gives $S_p = h_0^\intercal Z_p = (h^\intercal + \delta e_1^\intercal) Z_p \sim \mathcal{N}(\xi_p(h^\intercal + \delta e_1^\intercal)g, h_0^\intercal \Sigma_p h_0) = \mathcal{N}(\delta\xi_p, v_{S,p})$. And note $g = g_0 + \delta e_2$, so (15) implies $\mathrm{Var}(T_p) = g_0^\intercal \Sigma_p^{-1}\Sigma_p\Sigma_p^{-1} g_0/(g_0^\intercal \Sigma_p^{-1} g_0)^2 = (g_0^\intercal \Sigma_p^{-1} g_0)^{-1} = v_{T,p}$ and $\mathbb{E}(T_p) = g_0^\intercal \Sigma_p^{-1} g \xi_p / g_0^\intercal \Sigma_p^{-1} g_0 = g_0^\intercal \Sigma_p^{-1}(g_0 + \delta e_2)\xi_p / g_0^\intercal \Sigma_p^{-1} g_0 = (1 + \delta\theta_p)\xi_p$, so $T_p \sim \mathcal{N}((1 + \delta\theta_p)\xi_p, v_{T,p})$. Finally,

$$\mathrm{Cov}(S_p, T_p) = \mathrm{Cov}(h_0^\intercal Z_p, g_0^\intercal \Sigma_p^{-1} Z_p/(g_0^\intercal \Sigma_p^{-1} g_0)) = \frac{h_0^\intercal \Sigma_p \Sigma_p^{-1} g_0}{g_0^\intercal \Sigma_p^{-1} g_0} = \frac{h_0^\intercal g_0}{g_0^\intercal \Sigma_p^{-1} g_0} = 0.$$

So (17) is proved. Note the joint normality guarantees that zero correlation implies independence, and the rest of the statements easily follow.

Proof of Theorem 3

By Lemma 1 and (18), the log-likelihood of $\{(S_p^*, T_p^*) : p \in \mathcal{P}_{\mathfrak{a}_B}\}$ is

$$\ell\left(\delta, \{\xi_p^*\}_p\right) = -\frac{1}{2}\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \left[(S_p^* - \delta\lambda_p\xi_p^*)^2 + \{T_p^* - (1+\delta\theta_p)\xi_p^*\}^2\right] - \frac{|\mathcal{P}_{\mathfrak{a}_B}|}{2}\log(2\pi).$$

The score with respect to δ is

$$\frac{\partial}{\partial\delta}\ell\left(\delta, \{\xi_p^*\}_{p\in\mathcal{P}_{\mathfrak{a}_B}}\right) = \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \left[\lambda_p\xi_p^*(S_p^* - \delta\lambda_p\xi_p^*) + \theta_p\xi_p^*\{T_p^* - (1+\delta\theta_p)\xi_p^*\}\right]. \tag{33}$$

Under the restriction $\delta = 0$, the likelihood separates as

$$\ell\left(0, \{\xi_p^*\}_{p\in\mathcal{P}_{\mathfrak{a}_B}}\right) = -\frac{1}{2}\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\left[(S_p^*)^2 + (T_p^* - \xi_p^*)^2\right] - \frac{|\mathcal{P}_{\mathfrak{a}_B}|}{2}\log(2\pi),$$

so the restricted MLE is $\tilde{\xi}_p^* = T_p^*$. Let $\hat{\xi}_p^*(\delta)$ denote the profile MLE of ξ_p^* for a fixed δ where the profiled log-likelihood is

$$\ell_{\text{prof}}(\delta) = \ell\left(\delta, \{\hat{\xi}_p^*(\delta)\}_{p\in\mathcal{P}_{\mathfrak{a}_B}}\right).$$

By the chain rule, the profile score restricted at the null is

$$\ell'_{\text{prof}}(0) = \frac{\partial}{\partial\delta}\ell\Big|_{\delta=0,\,\xi_p^*=\tilde{\xi}_p^*} + \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\frac{\partial}{\partial\xi_p^*}\ell\left(0, \{\xi_p^*\}_{p\in\mathcal{P}_{\mathfrak{a}_B}}\right)\Big|_{\xi_p^*=\tilde{\xi}_p^*}\cdot\frac{\partial\hat{\xi}_p^*(0)}{\partial\delta}.$$

The second term is zero by the first-order condition of the restricted likelihood. Hence, substituting $\tilde{\xi}_p^* = T_p^*$ into $\ell'_{\text{prof}}(0)$ and noting (33) gives

$$\ell'_{\text{prof}}(0) = \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\left[\lambda_p T_p^* S_p^* + \theta_p T_p^*(T_p^* - T_p^*)\right] = \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\lambda_p T_p^* S_p^*.$$

Under H_0, conditional on $(T_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}}$, the conditional variance of the profile score is

$$\text{Var}\left(\ell'_{\text{prof}}(0) \mid (T_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}}\right) = \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\lambda_p^2(T_p^*)^2.$$

The conditional LM test statistic is therefore the squared profile score under the null hypothesis divided by its null conditional variance:

$$LM = \frac{\ell'_{\text{prof}}(0)^2}{\text{Var}\left(\ell'_{\text{prof}}(0) \mid (T_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}}\right)} = \frac{\left(\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\lambda_p S_p^* T_p^*\right)^2}{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}}\lambda_p^2(T_p^*)^2}.$$

Under H_0, $(S_p^*)_{p\in\mathcal{P}_{\mathfrak{a}_B}} \sim \mathcal{N}(0, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$ and $S_p^* \perp\!\!\!\perp T_p^*$ for all p. Therefore $LM \mid T^* \sim \chi_1^2$. Since the denominator is positive with probability one, the unconditional null distribution is also χ_1^2.

Proof of Theorem 4

In the limiting experiment $Z_p = (X_p, Y_p)^\intercal \sim \mathcal{N}(\xi_p g, \Sigma_p)$, transform Z_p into (S_p^*, T_p^*). By (18),

$$U_p \equiv (S_p^*, T_p^*)^\intercal = A_p Z_p \sim \mathcal{N}\left(\xi_p^* \begin{pmatrix} \delta\lambda_p \\ 1+\delta\theta_p \end{pmatrix}, I_2\right), \quad A_p = \begin{pmatrix} h_0^\intercal/\sqrt{v_{S,p}} \\ \sqrt{v_{T,p}} g_0^\intercal \Sigma_p^{-1} \end{pmatrix},$$

where $\delta, \xi_p^*, \lambda_p, \theta_p, v_{S,p}, v_{T,p}$ are defined in Section 5.1, and U_p's are independent across $p \in \mathcal{P}_{\mathfrak{a}_B}$. Therefore, the Gaussian limit experiment Z_p with parameters (β, ξ_p) can be transformed to $U_p = (S_p^*, T_p^*)^\intercal$ with parameters (δ, ξ_p^*) and the testing problem is equivalent to $H_0 : \delta = 0$ against $H_1 : \delta \neq 0$. Let $\mathfrak{m}_p(\delta) \equiv (\delta\lambda_p, 1+\delta\theta_p)^\intercal$. Then the log-likelihood of the $\{U_p\}_{p\in\mathcal{P}_{\mathfrak{a}_B}}$ is

$$\ell(\delta, \{\xi_p^*\}) = -\frac{1}{2}\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \|U_p - \xi_p^* \mathfrak{m}_p(\delta)\|_2^2 - \frac{|\mathcal{P}_{\mathfrak{a}_B}|}{2}\log(2\pi).$$

For any δ, profiling out ξ_p^* gives its profile MLE $\hat{\xi}_p^*(\delta) = \mathfrak{m}_p^\intercal(\delta) U_p / [\mathfrak{m}_p^\intercal(\delta)\mathfrak{m}_p(\delta)]$. The concentrated log-likelihood is therefore

$$\ell_c(\delta) = -\frac{1}{2}\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \|U_p\|_2^2 + \frac{1}{2}\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{[\mathfrak{m}_p^\intercal(\delta)U_p]^2}{\mathfrak{m}_p^\intercal(\delta)\mathfrak{m}_p(\delta)} - \frac{|\mathcal{P}_{\mathfrak{a}_B}|}{2}\log(2\pi).$$

Under the null $\delta = 0$, $\mathfrak{m}_p(0) = e_2$ and $[\mathfrak{m}_p^\intercal(0)U_p]^2/\mathfrak{m}_p^\intercal(0)\mathfrak{m}_p(0) = (T_p^*)^2$. So

$$\begin{aligned} LR &= \sup_{\delta\in\mathbb{R}} 2\{\ell_c(\delta) - \ell_c(0)\} = \sup_{\delta\in\mathbb{R}} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{[\mathfrak{m}_p^\intercal(\delta)U_p]^2}{\mathfrak{m}_p^\intercal(\delta)\mathfrak{m}_p(\delta)} - \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} (T_p^*)^2 \\ &= \sup_{\delta\in\mathbb{R}} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{\{\delta\lambda_p S_p^* + (1+\delta\theta_p)T_p^*\}^2}{\delta^2\lambda_p^2 + (1+\delta\theta_p)^2} - Q_T. \end{aligned}$$

Suppose $\Sigma_p = \Upsilon_p \bar{\Sigma}$. Let $\mathcal{O} \equiv \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$. So $\Sigma_p^{-1} = \mathcal{O}^\intercal \Sigma_p \mathcal{O}/\det\Sigma_p$, and $\mathcal{O}g_0 = h_0$, $\mathcal{O}e_2 = -e_1$. Hence $g_0^\intercal\Sigma_p^{-1}g_0 = v_{S,p}/\det\Sigma_p$, $g_0^\intercal\Sigma_p^{-1}e_2 = -h_0^\intercal\Sigma_p e_1/\det\Sigma_p$, and

$$v_{T,p} = \frac{\det\Sigma_p}{v_{S,p}}, \qquad \lambda_p = \frac{\sqrt{\det\Sigma_p}}{h_0^\intercal\Sigma_p h_0}, \qquad \theta_p = -\frac{h_0^\intercal\Sigma_p e_1}{h_0^\intercal\Sigma_p h_0}.$$

Therefore, given $\Sigma_p = \Upsilon_p \bar{\Sigma}$, λ_p and θ_p do not depend on p, implying that $\mathfrak{m}_p(\delta)$ is free of p and can be written as $\mathfrak{m}(\delta)$. By the Rayleigh quotient ratio,

$$\begin{aligned} LR = LR^h &= \sup_{\delta \in \mathbb{R}} \frac{\mathfrak{m}^\intercal(\delta) Q \mathfrak{m}(\delta)}{\mathfrak{m}^\intercal(\delta)\mathfrak{m}(\delta)} - Q_T = \lambda_{\max}(Q) - Q_T \\ &= \frac{1}{2}\left[(Q_S - Q_T) + \sqrt{(Q_S + Q_T)^2 - 4(Q_S Q_T - Q_{ST}^2)}\right]. \end{aligned}$$

So (20) and (21) are proved. For the rest proof, see Lemma 4 in the Supplement.

Proof of Theorem 5

(a) In the limit experiment, $S^* \perp\!\!\!\perp D$ because $D = \Lambda T^*$, $S^* \perp\!\!\!\perp T^*$ by Lemma 1, and $T^* \mapsto D$ is a bijective mapping. It follows that $S^* \mid D \sim \mathcal{N}(m, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$ where $m = \delta \Lambda \xi^*$. Let $P_D \equiv D(D^\intercal D)^{-1} D^\intercal$ and $M_D \equiv I - P_D$. Then by definition $LM = S^{*\intercal} P_D S^*$ and $J = S^{*\intercal} M_D S^*$. The matrices P_D and M_D are symmetric matrices with ranks 1 and $|\mathcal{P}_{\mathfrak{a}_B}| - 1$, respectively. Cochran's theorem gives

$$LM \mid D \sim \chi_1^2(m^\intercal P_D m), \qquad J \mid D \sim \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}(m^\intercal (I - P_D) m), \qquad LM \perp\!\!\!\perp J \mid D.$$

Note that under the null, $\delta = 0$ implies $m = 0$; thus

$$LM + a(D) \cdot J \mid D \overset{\mathcal{L}, H_0}{=} \chi_1^2 + a(D) \cdot \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1},$$

where $\chi_1^2 \perp\!\!\!\perp \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}$. Therefore $\mathbb{E}_{H_0}(\varphi^\alpha_{CLC, a(D)} \mid D) = \alpha$.
(b) In the Gaussian experiment, strong and semi-strong identification correspond to diverging rescaled jumps:

$$\xi^* = \xi_n^* = s_n u_n, \qquad s_n = \Delta_n^{r-1/2} \to \infty, \qquad u_n \to u \in \mathbb{R}_*^{|\mathcal{P}_{\mathfrak{a}_B}|},$$

where $r = 0$ represents the strong case and $0 < r < 1/2$ indicates the semi-strong case. Consider the null and local alternatives $\delta_n = b/s_n$, where $b \in \mathbb{R}$ is fixed and $b = 0$ gives the null. Then

$$\mu_D = \mu_{D,n} = s_n \Lambda u_n + b \Lambda \Theta u_n, \qquad \frac{\mu_{D,n}}{s_n} \to v \equiv \Lambda u.$$

Thus it suffices to show that $a_{\text{MMRU}}(\mu_{D,n}) \to 0$ if $\mu_{D,n}$ diverges. First, for every $C > 0$, let $B_C \equiv \{m \in \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} : \|m\| \leq C\}$. Then

$$d_H(\mathcal{M}_D(\mu_{D,n}) \cap B_C, \text{span}(v) \cap B_C) \to 0, \tag{34}$$

where $\mu_{D,n} = s_n v_n$, the sequence $v_n \to v$, and d_H denotes Hausdorff distance with the Euclidean metric $d(\cdot, \cdot)$:

$$d_H(A_1, A_2) = \max\{\sup_{x_1 \in A_1} d(x_1, A_2), \sup_{x_2 \in A_2} d(x_2, A_1)\}$$

To see this, first, let $m_n \in \mathcal{M}_D(\mu_{D,n}) \cap B_C$. Then $m_n = \delta_n \Lambda \xi_n^*$, so $\mu_{D,n} = \Lambda \xi_n^* + \delta_n \Lambda \Theta \xi_n^* = \Lambda \xi_n^* + \Theta m_n$ by (23), and therefore

$$m_n = \delta_n(\mu_{D,n} - \Theta m_n) = \delta_n(s_n v_n - \Theta m_n).$$

Choose a coordinate j with $v_j \neq 0$. The identity gives

$$\delta_n = \frac{(m_n)_j}{s_n (v_n)_j - \theta_j (m_n)_j} = O(s_n^{-1}), \qquad h_n \equiv s_n \delta_n = O(1),$$

noting $\|m_n\| \leq C$ for $m_n \in B_C$. So $m_n = h_n v_n - h_n \Theta m_n / s_n$. Consequently,

$$m_n = h_n (I + h_n \Theta / s_n)^{-1} v_n = h_n v + o(1), \tag{35}$$

so $d(m_n, \text{span}(v) \cap B_C) \to 0$ as n goes to infinity. Second, pick any $hv \in \text{span}(v) \cap B_C$ where h is bounded. Note by (35), there exists an $m_n' \in \mathcal{M}_D(\mu_D)$ s.t. $m_n' = h(I + h\Theta/s_n)^{-1} v_n = hv + o(1)$ with $\|m_n'\| < C$ if n is large. So $d(hv, \mathcal{M}_D(\mu_{D,n}) \cap B_C) \to 0$ as n goes to infinity. Therefore, (34) is proved. According to Theorem 5 of Andrews (2016b), $a_{\text{MMRU}} \to 0$.

References

Acosta, M., C. M. Brennan, and M. M. Jacobson (2024): "Constructing high-frequency monetary policy surprises from SOFR futures," *Economics Letters*, 242, 111873.

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang (2026): "High-Frequency In-

struments With Time-Varying Reliability: Understanding Identification in Macroeconomics," *Journal of Applied Econometrics*, 41(5), 539–548.

Anderson, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2003): "Micro Effects of Macro Announcements: Real-time Price Discovery in Foreign Exchange," *American Economic Review*, 93(1), 38–62.

Anderson, T. W., and H. Rubin (1949): "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, 20(1), 46–63.

Andrews, D. W., and X. Cheng (2012): "Estimation and Inference with Weak, Semi-strong, and Strong Identification," *Econometrica*, 80(5), 2153–2211.

Andrews, D. W., M. J. Moreira, and J. H. Stock (2006): "Optimal Two-sided Invariant Similar Tests for Instrumental Variables Regression," *Econometrica*, 74(3), 715–752.

Andrews, I. (2016a): "Conditional Linear Combination Tests for Weakly Identified Models," *Econometrica*, 84(6), 2155–2182.

——— (2016b): "Supplement to Conditional Linear Combination Tests for Weakly Identified Models: Appendix," *Econometrica Supplementary Material*, 84(6), 1–17.

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

Bauer, M. D., and E. T. Swanson (2023a): "An Alternative Explanation for the "Fed Information Effect"," *American Economic Review*, 113(3), 664–700.

——— (2023b): "A Reassessment of Monetary Policy Surprises and High-Frequency Identification," *NBER Macroeconomics Annual*, 37(1), 87–155.

Bollerslev, T., J. Li, and Y. Ren (2024): "Optimal Inference for Spot Regressions," *American Economic Review*, 114(3), 678–708.

Bollerslev, T., J. Li, and Y. Xue (2018): "Volume, Volatility, and Public News Announcements," *The Review of Economic Studies*, 85(4), 2005–2041.

Bollerslev, T., and V. Todorov (2011): "Estimation of Jump Tails," *Econometrica*, 79(6), 1727–1783.

Carpentier, A., C. Duval, and E. Mariucci (2021): "Total Variation Distance for Discretely Observed Lévy Processes: A Gaussian Approximation of the Small Jumps," *Annales de l'Institut Henri Poincaré, Probabilités et Statistiques*, 57(2), 901–939.

Casini, A., and A. McCloskey (2026): "Identification and Estimation of Causal Effects in High-Frequency Event Studies," .

Christensen, K., A. Timmermann, and B. Veliyev (2025): "Warp Speed Price Moves: Jumps after Earnings Announcements," *Journal of Financial Economics*, 167, 104010.

Dufour, J.-M. (1997): "Some Impossibility Theorems in Econometrics with Applications to Structural and Dynamic Models," *Econometrica: Journal of the Econometric Society*, pp. 1365–1387.

Feir, D., T. Lemieux, and V. Marmer (2016): "Weak Identification in Fuzzy Regression Discontinuity Designs," *Journal of Business and Economic Statistics*, 34(2), 185–196.

Golez, B., and B. Matthies (2025): "Fed Information Effects: Evidence from the Equity Term Structure," *Journal of Financial Economics*, 165, 103988.

Gürkaynak, R. S., B. Sack, and E. T. Swanson (2005): "Do Actions Speak Louder Than Words? The Response of Asset Prices to Monetary Policy Actions and Statements," *International Journal of Central Banking*.

Gürkaynak, R. S., E. T. Swanson, and J. H. Wright (2026): "High-Frequency Identification," Working paper, August 2026.

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

Jacod, J., and V. Todorov (2010): "Do Price and Volatility Jump Together?," *The Annals of Applied Probability*, 20(4), 1425–1469.

Känzig, D. R. (2021): "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements," *American Economic Review*, 111(4), 1092–1125.

Kleibergen, F. (2002): "Pivotal Statistics for Testing Structural Parameters in Instrumental Variables Regression," *Econometrica*, 70(5), 1781–1803.

——— (2005): "Testing Parameters in GMM Without Assuming That They Are Identified," *Econometrica*, 73(4), 1103–1123.

Kuttner, K. N. (2001): "Monetary Policy Surprises and Interest Rates: Evidence from the Fed Funds Futures Market," *Journal of Monetary Economics*, 47(3), 523–544.

Lahaye, J., S. Laurent, and C. J. Neely (2011): "Jumps, Cojumps and Macro announcements," *Journal of Applied Econometrics*, 26(6), 893–921.

Lehmann, E. L., and J. P. Romano (2005): *Testing Statistical Hypotheses*. Springer, 3rd edn.

Li, J. (2013): "Robust Estimation and Inference for Jumps in Noisy High Frequency Data: A Local-to-Continuity Theory for the Pre-Averaging Method," *Econometrica*, 81(4), 1673–1693.

Li, J., V. Todorov, and G. Tauchen (2017): "Jump Regressions," *Econometrica*, 85(1), 173–195.

Maheu, J. M., and T. H. McCurdy (2004): "News Arrival, Jump dynamics, and Volatility Components for Individual Stock Returns," *The Journal of Finance*, 59(2), 755–793.

Mancini, C. (2001): "Disentangling the Jumps of the Diffusion in a Geometric Jumping Brownian Motion," *Giornale dell'Istituto Italiano degli Attuari*, 64, 19–47.

Mikusheva, A. (2010): "Robust confidence sets in the presence of weak instruments," *Journal of Econometrics*, 157(2), 236–247.

Moreira, H., and M. J. Moreira (2019): "Optimal Two-Sided Tests for Instrumental Variables Regression with Heteroskedastic and Autocorrelated Errors," *Journal of Econometrics*, 213(2), 398–433.

Moreira, M. J. (2003): "A Conditional Likelihood Ratio Test for Structural Models," *Econometrica*, 71(4), 1027–1048.

——— (2009): "Tests with Correct Size When Instruments Can Be Arbitrarily Weak," *Journal of Econometrics*, 152(2), 131–140.

Müller, U. K. (2011): "Efficient Tests under a Weak Convergence Assumption," *Econometrica*, 79(2), 395–435.

Nakamura, E., and J. Steinsson (2018): "High-Frequency Identification of Monetary Non-neutrality: the Information Effect," *The Quarterly Journal of Economics*, 133(3), 1283–1330.

Ramey, V. A. (2016): "Macroeconomic Shocks and Their Propagation," *Handbook of macroeconomics*, 2, 71–162.

Staiger, D., and J. Stock (1997): "Instrumental Variables Regression with Weak Instruments," *Econometrica*, 65(3), 557–586.

Stock, J. H., and M. W. Watson (2012): "Disentangling the Channels of the 2007-09 Recession.," *Brookings Papers on Economic Activity*.

Todorov, V., and T. Bollerslev (2010): "Jumps and Betas: A New Framework for Disentangling and Estimating Systematic Risks," *Journal of Econometrics*, 157(2), 220–235.

Van der Vaart, A. W. (2000): *Asymptotic Statistics*. Cambridge University Press.

Supplement to "Weak Identification in News-Jump Regressions"

Jiezheng Wei*

October 29, 2026

Abstract

This supplement provides the proofs of lemmas, auxiliary theoretical results, additional simulation results, and further empirical findings of Wei (2026), "Weak Identification in News-Jump Regressions".

S1. Proof of Technical Lemmas and the Corollary

Lemma 2. Let $z, t, \lambda, \theta \in \mathbb{R}^P$ where P is an arbitrary finite integer and $\lambda_p > 0$ for every p. Define a function by taking the supremum

$$L(z; t, \lambda, \theta) \equiv \sup_{\delta \in \mathbb{R}} \left\{ \sum_{p=1}^{P} \frac{[\delta \lambda_p z_p + (1 + \delta\theta_p) t_p]^2}{\delta^2 \lambda_p^2 + (1 + \delta\theta_p)^2} - \|t\|^2 \right\}. \tag{S1}$$

Then $0 \leq L(z; t, \lambda, \theta) \leq \|z\|^2$.

Proof of Lemma 2. For each p and $\delta \in \mathbb{R}$, define

$$\mathfrak{m}_p(\delta) = \big(\delta\lambda_p,\, 1 + \delta\theta_p\big)^\intercal \quad \text{and} \quad \mathfrak{n}_p = (z_p, t_p)^\intercal.$$

Suppose $\mathfrak{m}_p(\delta) = 0$. Since $\lambda_p > 0$, the first coordinate of $\mathfrak{m}_p(\delta)$ would imply $\delta\lambda_p = 0$, and therefore $\delta = 0$; but its second coordinate would then equal $1 + \delta\theta_p = 1$, a contradiction. Thus $\mathfrak{m}_p(\delta) \neq 0$ for every $\delta \in \mathbb{R}$, and in particular $\|\mathfrak{m}_p(\delta)\|^2 > 0$. By the Cauchy–Schwarz inequality,

$$0 \leq \frac{\big[\mathfrak{m}_p(\delta)^\intercal \mathfrak{n}_p\big]^2}{\|\mathfrak{m}_p(\delta)\|^2} \leq \|\mathfrak{n}_p\|^2 = z_p^2 + t_p^2.$$

*School of Economics, Singapore Management University; E-mail: jzwei.2022@smu.edu.sg.

Consequently, for every $\delta \in \mathbb{R}$,

$$\sum_{p=1}^{P} \frac{[\delta\lambda_p z_p + (1+\delta\theta_p)t_p]^2}{\delta^2\lambda_p^2 + (1+\delta\theta_p)^2} - \|t\|^2 \leq \sum_{p=1}^{P}(z_p^2 + t_p^2) - \|t\|^2 = \|z\|^2.$$

Taking the supremum over δ establishes the upper bound. For the lower bound, evaluating the expression at $\delta = 0$ gives $\sum_{p=1}^{P} t_p^2 - \|t\|^2 = 0$. Hence $0 \leq L(z; t, \lambda, \theta) \leq \|z\|^2$ as claimed. □

Lemma 3. The limit $\tilde{b}$ defined in (30) is well defined and bounded in probability.

Proof of Lemma 3. Write

$$\mathcal{H}_p \equiv e_1^\intercal R_p + \xi_p, \qquad \mathfrak{D} \equiv \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \mathcal{H}_p^2.$$

First, $\mathfrak{D} > 0$ almost surely. On $\Omega(\mathfrak{a}_B)$, Assumption 2 implies that $\xi_p \neq 0$ for at least one $p \in \mathcal{P}_{\mathfrak{a}_B}$. Conditional on the sigma-field $\mathcal{F}$ and κ_p, it follows that

$$e_1^\intercal R_p \mid \mathcal{F}, \kappa_p \sim \mathcal{N}(0, \mathbb{V}_p), \qquad \mathbb{V}_p \equiv \kappa_p e_1^\intercal c_{\mathfrak{a}_p -} e_1 + (1-\kappa_p) e_1^\intercal c_{\mathfrak{a}_p} e_1 \geq 0.$$

If $\mathbb{V}_p > 0$, this conditional distribution is atomless for being normal, and hence $\mathbb{P}(\mathcal{H}_p = 0 \mid \mathcal{F}, \kappa_p) = 0$. If $\mathbb{V}_p = 0$, then $e_1^\intercal R_p = 0$ almost surely, implying $\mathcal{H}_p = \xi_p \neq 0$ whenever $\xi_p \neq 0$. Thus it follows that, noting $|\mathcal{P}_{\mathfrak{a}_B}| < \infty$,

$$\begin{aligned}\mathbb{P}(\mathfrak{D} = 0, \Omega(\mathfrak{a}_B)) &= \mathbb{P}\left(\{\mathcal{H}_p = 0, \forall p \in \mathcal{P}_{\mathfrak{a}_B}\} \bigcap \left(\cup_{p \in \mathcal{P}_{\mathfrak{a}_B}} \{\xi_p \neq 0\}\right)\right) \\ &\leq \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \mathbb{P}(\mathcal{H}_p = 0, \xi_p \neq 0) = 0.\end{aligned}$$

Hence $\mathfrak{D} > 0$ almost surely on $\Omega(\mathfrak{a}_B)$. This also verifies that the singular set $\{\mathfrak{D} = 0\}$ of the ratio map used in the continuous mapping theorem has probability zero.

Finally, let $\mathcal{H} = (\mathcal{H}_p)_{p \in \mathcal{P}_{\mathfrak{a}_B}}$ and $\varsigma(\beta) \equiv (\varsigma_p(\beta))_{p \in \mathcal{P}_{\mathfrak{a}_B}}$. The number of announcements is finite and every component of $\mathcal{H}$ and $\varsigma(\beta)$ is finite almost surely. By the Cauchy–Schwarz inequality,

$$|\tilde{b}| = \frac{|\mathcal{H}^\intercal \varsigma(\beta)|}{\|\mathcal{H}\|^2} \leq \frac{\|\varsigma(\beta)\|}{\|\mathcal{H}\|}.$$

For any $\varepsilon > 0$, since $\mathbb{P}(\|\mathcal{H}\| = 0) = 0$, choose $\delta > 0$ such that $\mathbb{P}(\|\mathcal{H}\| \leq \delta) < \varepsilon/2$, while the almost sure finiteness of $\|\varsigma(\beta)\|$ permits a choice of $K < \infty$ such that $\mathbb{P}(\|\varsigma(\beta)\| > K) < \varepsilon/2$. Additionally,

$$\mathbb{P}\left(|\tilde{b}| > K/\delta\right) \leq \mathbb{P}(\|\mathcal{H}\| \leq \delta) + \mathbb{P}(\|\varsigma(\beta)\| > K) < \varepsilon.$$

Thus $\tilde{b}$ is a well-defined random variable and $\tilde{b} = O_{\mathbb{P}}(1)$.

□

Lemma 4. The conditional random variable which is equal in law to LR in the limit experiment in Theorem 4(a) is well defined, and both of the CLR and CLR^h tests are conditionally similar.

Proof. The first claim is that the conditional random variable in Theorem 4(a) is well defined and continuously distributed. Define a compact set $\mathbb{S}^1_+ \equiv \{(a,b) \in \mathbb{R}^2 : a^2 + b^2 = 1,\ b \geq 0\}$. Relabel the indices in $\mathcal{P}_{\mathfrak{a}_B}$ by $\{1, 2, \cdots, P\}$. For $(a,b) \in \mathbb{S}^1_+$, define

$$G_{a,b}(z,t) \equiv \sum_{p=1}^{P} \frac{\{a\lambda_p z_p + (b + a\theta_p)t_p\}^2}{a^2\lambda_p^2 + (b + a\theta_p)^2} - \|t\|^2.$$

Every denominator is strictly positive for $(a,b) \in \mathbb{S}^1_+$. For $b > 0$, $G_{a,b}$ equals the objective defining L in (S1) at $\delta = a/b$, whereas the two points with $b = 0$ give the common limit as $\delta \to \pm\infty$. Hence

$$L(z; t, \lambda, \theta) = \max_{(a,b)\in\mathbb{S}^1_+} G_{a,b}(z,t). \tag{S2}$$

Since $G_{a,b}(z,t)$ is jointly continuous and $\mathbb{S}^1_+$ is compact, the maximum in (S2) is finite, and by Berge's maximum theorem (see Theorem 17.31 of Aliprantis and Border (2006)), $L(z; t, \lambda, \theta)$ is jointly continuous in (z, t, λ, θ) on $\mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \times \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \times (0,\infty)^{|\mathcal{P}_{\mathfrak{a}_B}|} \times \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|}$, and therefore Borel measurable. Lemma 2 additionally gives $0 \leq L(z; t, \lambda, \theta) \leq \|z\|^2$.

The next claim is that, for each fixed t, every level set of $z \mapsto L(z; t, \lambda, \theta)$ is Lebesgue-null. Consider first a level $c > 0$. Holding $z_{-1} = (z_2, \ldots, z_P)$ fixed, suppose that three points $r_1 < r_2 < r_3$ satisfy $L((r_j, z_{-1}); t, \lambda, \theta) = c$ for $j = 1, 2, 3$. Choose

a maximizer (a, b) in (S2) at r_2. Since $G_{0,1} = 0$, necessarily $a \neq 0$. Moreover, for $z = (r, z_{-1})$, note $\lambda_1 > 0$, so

$$\frac{\partial^2}{\partial r^2} G_{a,b}((r, z_{-1}), t) = \frac{2a^2\lambda_1^2}{a^2\lambda_1^2 + (b + a\theta_1)^2} > 0.$$

Thus, writing $r_2 = sr_1 + (1 - s)r_3$ for some $s \in (0, 1)$, strict convexity gives $c = G_{a,b}((r_2, z_{-1}), t) < sG_{a,b}((r_1, z_{-1}), t) + (1 - s)G_{a,b}((r_3, z_{-1}), t) = c$, a contradiction. Every section parallel to the first coordinate axis therefore contains at most two points, so Fubini's theorem implies that the level set at c is Lebesgue-null:

$$\mathbb{P}(\{L(z; t, \lambda, \theta) = c\}) = \sum_{z_{-1}} \mathbb{P}(\{L((z_1, z_{-1}); t, \lambda, \theta) = c\} | z_{-1}) = 0$$

It remains to consider the level $c = 0$. Denote the objective in (S1) before taking the supremum by $R_\delta(z, t)$. Direct differentiation at the null gives

$$\frac{\partial}{\partial \delta} R_\delta(z, t)\bigg|_{\delta=0} = 2 \sum_{p=1}^{P} \lambda_p t_p z_p.$$

If $t \neq 0$ and $L(z; t, \lambda, \theta) = 0$, then the supremum implies $R_\delta(z, t) \leq 0 = R_0(z, t)$ for every δ, so the derivative above must vanish. Consequently,

$$\{z : L(z; t, \lambda, \theta) = 0\} \subseteq \left\{ z : \sum_{p=1}^{P} \lambda_p t_p z_p = 0 \right\},$$

which is a Lebesgue-null hyperplane because $\lambda_p > 0$ for every p and $t \neq 0$. If $t = 0$, then for every $z \neq 0$ and every $\delta \neq 0$,

$$R_\delta(z, 0) = \sum_{p=1}^{P} \frac{\delta^2 \lambda_p^2 z_p^2}{\delta^2 \lambda_p^2 + (1 + \delta\theta_p)^2} > 0,$$

and hence the zero-level set is the singleton $\{0\}$. Negative level sets are empty by Lemma 2. This proves that $L(\zeta; t, \lambda, \theta)$ is atomless for every fixed t whenever $\zeta \sim \mathcal{N}(0, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$. Its distribution has no gaps either. Indeed, Lemma 2 gives $L(0; t, \lambda, \theta) = 0$, while $L(re_1; t, \lambda, \theta) \geq G_{1,0}(re_1, t) \to \infty$ as $|r| \to \infty$. Since $\mathbb{R}^P$ is connected and L is continuous, the range of the random variable $L(\zeta, t, \lambda, \theta)$ is therefore $[0, \infty)$. For every $0 \leq x < y$, the set $\{z : x < L(z; t, \lambda, \theta) < y\}$ is nonempty and open, and hence

has positive probability under the standard multivariate Gaussian distribution. Thus the distribution function of $L(\zeta; t, \lambda, \theta)$ is continuous and strictly increasing on $[0, \infty)$.

To identify the conditional distribution formally, let γ_P be the standard P-dimensional Gaussian measure on $\mathbb{R}^P$ and define, for every Borel set $A \subseteq \mathbb{R}$,

$$\mathcal{K}_{\lambda,\theta}(t, A) \equiv \int_{\mathbb{R}^P} 1\{L(z; t, \lambda, \theta) \in A\}\, \gamma_P(\mathrm{d}z).$$

Joint measurability of L makes $\mathcal{K}_{\lambda,\theta}$ a probability kernel. Under H_0, Lemma 1 gives $S^* \sim \mathcal{N}(0, I_P)$ and $S^* \perp\!\!\!\perp T^*$. Therefore, for all Borel sets $A \subseteq \mathbb{R}$ and $B \subseteq \mathbb{R}^P$,

$$\mathbb{P}\{LR \in A,\ T^* \in B\} = \int_B \mathcal{K}_{\lambda,\theta}(t, A)\, \mathbb{P}_{T^*}(\mathrm{d}t),$$

which verifies that $\mathcal{K}_{\lambda,\theta}(T^*, \cdot)$ is a regular conditional distribution of LR given T^*. Equivalently,

$$LR \mid T^* = t \overset{\mathcal{L}}{=} L(\zeta; t, \lambda, \theta), \qquad \zeta \sim \mathcal{N}(0, I_P),$$

which is the distribution F_{T^*} in Theorem 4. Let $F_{\lambda,\theta}(t, x) \equiv \mathcal{K}_{\lambda,\theta}(t, (-\infty, x])$ and

$$cv^{\alpha}_{CLR}(t) \equiv \inf\{x : F_{\lambda,\theta}(t, x) \geq 1 - \alpha\}.$$

The critical-value map is jointly Borel measurable in (t, λ, θ), and the preceding argument makes $F_{\lambda,\theta}(t, \cdot)$ continuous and strictly increasing on its support. It follows that

$$\mathbb{P}_{H_0}\{LR > cv^{\alpha}_{CLR}(t) \mid T^* = t\} = \alpha.$$

Thus, for the rejection rule in the theorem given by $\varphi^{\alpha}_{CLR} = 1\{LR > cv^{\alpha}_{CLR}(T^*)\}$, it follows that $\mathbb{E}_{H_0,\mathbb{P}}[\varphi^{\alpha}_{CLR} \mid T^*] = \alpha$.

Finally, consider the conditional similarity of the CLR^h test in part (b). Under H_0, fix $T^* = t$ and write $q = t^{\intercal}t$ which is a realization of Q_T. If $q > 0$, let $u = t/\sqrt{q}$ and define

$$K = (u^{\intercal}S^*)^2, \qquad J = S^{*\intercal}(I_{|\mathcal{P}_{\mathfrak{a}_B}|} - uu^{\intercal})S^*.$$

By Cochran's theorem (Theorem 1.5 of Shao (2003)), $K \sim \chi^2_1, J \sim \chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}, K \perp\!\!\!\perp J$.

Furthermore, $Q_S = K + J, Q_{ST}^2 = qK$, and $Q_S Q_T - Q_{ST}^2 = (K+J)q - qK = qJ$. Substitution into (21) yields

$$LR^h \mid T^* = t \overset{\mathcal{L}}{=} \frac{1}{2}\left\{K + J - q + \sqrt{(K+J+q)^2 - 4qJ}\right\}.$$

This proves (22): all occurrences of $\chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|-1}$ therein represent the same random variable J, χ^2_1 represents the random variable K, and $J \perp\!\!\!\perp K$. And this conditional distribution is atomless: for $q = 0$ it is $\chi^2_{|\mathcal{P}_{\mathfrak{a}_B}|}$, whereas for $q > 0$ the displayed expression is a strictly increasing function of the continuously distributed random variables J and K. Thus, with $\varphi^{\alpha}_{CLR^h} = 1\{LR^h \mid T^* > cv^{\alpha}_{CLR^h}(T^*)\}$, the same probability-kernel and quantile argument as in part (a) yields $\mathbb{E}_{H_0,\mathbb{P}}[\varphi^{\alpha}_{CLR^h} \mid T^*] = \alpha$ a.s., which proves the conditional similarity. □

Proof of Corollary 1

Let $m = \mathbb{E}(S^*) = \delta\Lambda\xi^*$, so $S^* \sim \mathcal{N}(m, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$. By (18), the transformed Gaussian experiment can be expressed as

$$\begin{pmatrix} S^* \\ T^* \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} m \\ (I_{|\mathcal{P}_{\mathfrak{a}_B}|} + \delta\Theta)\xi^* \end{pmatrix}, I_{2|\mathcal{P}_{\mathfrak{a}_B}|}\right).$$

And the testing problem becomes

$$H_0 : m = 0 \qquad \text{against} \qquad H_1 : m \in \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \setminus \{0\}$$

Consequently, for every fixed t, $S^* \mid T^* = t \sim \mathcal{N}(m, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$. Since Λ is nonsingular and $\xi^* \in \Xi = \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \setminus \{0\}$, the set $\{m = \delta\Lambda\xi^* : \delta \neq 0,\ \xi^* \in \Xi\} = \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|} \setminus \{0\}$. Thus, conditional on $T^* = t$, the weak-jump experiment is precisely the Gaussian mean problem stated in the Corollary.

Next, consider the geometry of the two acceptance regions. First consider the CLR^h test. For $t \neq 0$, define

$$K(s,t) \equiv \frac{(t^\intercal s)^2}{t^\intercal t} \geq 0, \qquad J(s,t) \equiv s^\intercal s - K(s,t) \geq 0, \qquad q \equiv t^\intercal t \geq 0,$$

where $J \geq 0$ is by Cauchy–Schwarz. And the test $LR^h = LR^h(S^*, T^*)$ is a function of S^* and T^*. Given $T^* = t$ for each fixed $t \in \mathbb{R}^{|\mathcal{P}_{\mathfrak{a}_B}|}$, $Q_T = q$, $Q_S = K + J$, and

$Q_{ST}^2 = (t^\intercal s)^2 = (t^\intercal t) \cdot \frac{(t^\intercal s)^2}{t^\intercal t} = qK$. By (21),

$$\begin{aligned} LR^h &= \frac{1}{2}\left\{K + J - q + \sqrt{(K+J+q)^2 - 4[(K+J)q - qK]}\right\} \\ &= \frac{1}{2}\left\{K + J - q + \sqrt{(K+J-q)^2 + 4qK}\right\}. \end{aligned}$$

Note LR^h is the nonnegative root of the quadratic equation in ℓ, $\ell^2 - (K+J-q)\ell - Kq = 0$. Let $c_h^\alpha(t)$ denote the $(1-\alpha)$-quantile of the null distribution of $LR^h(S^*, t)$, where $S^* \sim \mathcal{N}(0, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$. The critical value satisfies $c_h^\alpha(t) > 0$. So

$$\begin{aligned} LR^h(S^*,t) \le c_h^\alpha(t) &\iff c_h^\alpha(t)^2 - (K+J-q)c_h^\alpha(t) - Kq \ge 0 \\ &\iff c_h^\alpha(t)(c_h^\alpha(t)+q) - K(c_h^\alpha(t)+q) - Jc_h^\alpha(t) \ge 0 \\ &\overset{(\star)}{\iff} \frac{K(S^*,t)}{c_h^\alpha(t)} + \frac{J(S^*,t)}{c_h^\alpha(t)+q} \le 1, \end{aligned}$$

where step $(\star)$ is due to $c_h^\alpha(t)(c_h^\alpha(t)+q) > 0$. Moreover, if s is decomposed into a component parallel to t and a component orthogonal to t, namely

$$s = s_{\mathfrak{p}} + s_\perp, \qquad \text{where } s_{\mathfrak{p}} \equiv u^\intercal s u, \quad u \equiv t/\sqrt{t^\intercal t},$$

then $K(s,t) = \|s_{\mathfrak{p}}\|^2$ and $J(s,t) = \|s_\perp\|^2$. So the acceptance region of the CLR^h test conditional on $T^* = t$ is as follows

$$\mathcal{A}_{CLR^h}^{1-\alpha}(t) = \left\{s : \frac{\|s_{\mathfrak{p}}\|^2}{c_h^\alpha(t)} + \frac{\|s_\perp\|^2}{c_h^\alpha(t)+q} \le 1\right\}$$

which is a closed, bounded ellipsoid with nonempty interior in the transformed coordinates. If $t = 0$, then $LR^h(S^*, 0) = (S^*)^\intercal S^*$, so in this case its acceptance region becomes a closed ball. In either case the acceptance region is convex and its boundary is a Lebesgue-null set.

For the CLR test under the general heterogeneity case, define

$$F_\delta(s,t) \equiv \begin{cases} \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{\{\delta\lambda_p s_p + (1+\delta\theta_p)t_p\}^2}{\delta^2\lambda_p^2 + (1+\delta\theta_p)^2} & \delta \in \mathbb{R} \\ \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{(\lambda_p s_p + \theta_p t_p)^2}{\lambda_p^2 + \theta_p^2} & \delta = \pm\infty \end{cases}$$

If $c^\alpha(t) > 0$ is the critical value used by φ^α_{CLR} given $T^* = t$, then

$$LR(s,t) \le c^\alpha(t) \iff \sup_\delta F_\delta(s,t) - t^\intercal t \le c^\alpha(t) \iff F_\delta(s,t) \le t^\intercal t + c^\alpha(t), \quad \forall \delta$$

So its acceptance region is

$$\mathcal{A}^{1-\alpha}_{CLR}(t) = \bigcap_{\delta \in \overline{\mathbb{R}}} \{s : F_\delta(s,t) \le t^\intercal t + c^\alpha(t)\}, \tag{S3}$$

where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}$. Note that, for each $\delta \in \overline{\mathbb{R}}$, the function $F_\delta(s,t)$ is a quadratic form in s so it is a continuous and convex function, which means each sublevel set member of the intersection (S3) is convex and closed. Therefore $\mathcal{A}^{1-\alpha}_{CLR}(t)$ is closed and convex. Moreover, when $\delta = \pm\infty$, note

$$F_{\pm\infty}(s,t) = \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{\lambda_p^2}{\lambda_p^2 + \theta_p^2} \left(s_p + \frac{\theta_p t_p}{\lambda_p} \right)^2 \to \infty, \quad \text{as } \|s\| \to \infty$$

and that $F_{\pm\infty}(s,t)$ is continuous in s. Therefore, the set $\{s : F_{-\infty}(s,t) \le t^\intercal t + c^\alpha(t)\} \cap \{s : F_{+\infty}(s,t) \le t^\intercal t + c^\alpha(t)\}$ is bounded. So $\mathcal{A}^{1-\alpha}_{CLR}(t)$ is bounded. Moreover, $F_\delta(0,t) \le t^\intercal t$ for every δ, with equality at $\delta = 0$, and hence $LR(0,t) = 0$. And according to the proof of Theorem 4, $\sup_\delta F_\delta(s,t)$ is continuous in s. It follows that $\mathcal{A}_{CLR}(t)$ has nonempty interior for containing $\{s = 0\}$, and its boundary is Lebesgue-null by the continuity of $\sup_\delta F_\delta(s,t)$ in s and the convexity of $\mathcal{A}^{1-\alpha}_{CLR}(t)$.

It remains to invoke the complete class result in Birnbaum (1955) in which the following exponential family, with a k-dimensional parameter ψ, is considered

$$f_\psi(s) = \exp\left\{ \psi_0 + \sum_{i=1}^k \psi_i t_i(s) + t_0(s) \right\} \tag{S4}$$

The random vector $S^* \sim \mathcal{N}(m, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$ has density $f_m(s)$:

$$f_m(s) = \exp\left\{ -\frac{1}{2}\|m\|^2 + m^\intercal s - \frac{1}{2}\|s\|^2 - \frac{|\mathcal{P}_{\mathfrak{a}_B}|}{2}\log(2\pi) \right\}.$$

Therefore, linking to (S4), $t_0(s) = -\|s\|^2/2 - (|\mathcal{P}_{\mathfrak{a}_B}|/2)\log(2\pi)$ is continuous, which verifies Birnbaum's Assumption 2. Let S denote the support of S^*, let

$$S' = \{s : a^\intercal s + a_0 \le 0\}, \qquad \|a\| = 1,$$

be any half-space, and let $s_0 \notin S'$. Choose a bounded neighborhood U of s_0 such that $\inf_{s\in U}(a^\intercal s + a_0) = \epsilon > 0$. For $r > 0$, set $m_r = ra$. Under $\mathcal{N}(m_r, I_{|\mathcal{P}_{\mathfrak{a}_B}|})$, $\mathbb{P}_{m_r}(S') = \Phi(-r - a_0)$. For all sufficiently large r, the Mills bound $\Phi(-x) \le \phi(x)/x$ therefore yields, uniformly over $s \in U$,

$$\begin{aligned}\frac{\phi_P(s - m_r)}{\mathbb{P}_{m_r}(S')} &\ge (2\pi)^{-(|\mathcal{P}_{\mathfrak{a}_B}|-1)/2}(r + a_0)\exp\left\{-\frac{\|s\|^2}{2} + \frac{a_0^2}{2} + r(a^\intercal s + a_0)\right\} \\ &\ge C_U(r + a_0)e^{r\epsilon} \longrightarrow \infty,\end{aligned}$$

where $C_U > 0$ is independent of r. This verifies Birnbaum's Assumption 3. By Theorem 4 of Birnbaum (1955), tests that are almost-everywhere equivalent to tests with open convex acceptance regions form a minimal complete class. Since both $\mathcal{A}^{1-\alpha}_{CLR}$ and $\mathcal{A}^{1-\alpha}_{CLR^h}$ differ from their open interiors only on a Lebesgue-null boundary, both conditional tests belong to this minimal essentially complete class. Thus φ^α_{CLR} and $\varphi^\alpha_{CLR^h}$ are admissible in every fixed-t conditional experiment by minimality.

S2. Justification of the Limit Problem

The following proposition supplies the link between the finite-sample testing procedures and the Gaussian limit experiment.

Proposition 1. Fix $\beta_0 \in \mathbb{R}$ and $|\mathcal{P}_{\mathfrak{a}_B}| < \infty$. Suppose that Assumptions 2, 1′, and 3 hold, that $c_{\mathfrak{a}_p}$ is positive definite for each $p \in \mathcal{P}_{\mathfrak{a}_B}$, and that $\hat{c}_{\mathfrak{a}_p}$ is a consistent estimator for $c_{\mathfrak{a}_p}$. Then

$$\left(S^*_n(\beta_0), T^*_n(\beta_0), \hat{\Lambda}_n(\beta_0), \hat{\Theta}_n(\beta_0)\right) \xrightarrow{\mathcal{L}\text{-}s} (S^*, T^*, \Lambda, \Theta)\,.$$

Suppose additionally that the plug-in conditional distribution functions are continuous and strictly increasing at their $(1-\alpha)$-quantiles almost surely, the numerical maximization error in (28) is $o_{\mathbb{P}}(1)$, and the number of conditional Monte Carlo draws satisfies $M \to \infty$. For the CLC test, suppose also that

$$\left(S^*_n(\beta_0), D_n(\beta_0), \hat{a}_{n,i}(\beta_0), \hat{\Lambda}_n(\beta_0), \hat{\Theta}_n(\beta_0)\right) \xrightarrow{\mathcal{L}} (S^*, D, a_i(D, \Lambda, \Theta), \Lambda, \Theta)\,, \quad D = \Lambda T^*,$$

where a_i is a measurable $[0,1]$-valued function. Under $H_0 : \beta = \beta_0$, the feasible LM, CLR, CLR^h, and CLC tests in Algorithm 2 satisfy, respectively,

$$\mathbb{P}\{\text{reject } H_0\} \longrightarrow \alpha.$$

Consequently, inversion of any one of these tests yields pointwise asymptotic coverage $1-\alpha$ at the true value of β.

Proof. The proof is immediate because the conditioning statistics are minimal sufficient for the jumps and complete, yielding a Neyman structure, which means conditional similarity. Hence weak convergence and the continuous mapping theorem imply equality between the power functions in the limit problem and the limits of the power functions in the finite-sample problem. □

S3. Additional Results under Strong Identification

This section presents additional results under strong identification. The first result establishes that, under additional regularity conditions, the LM and LTT-GLS tests are asymptotically equivalent against local alternatives of order $1/\sqrt{n}$. The second result shows that the LM and CLR tests are asymptotically equivalent and that the LM test is uniformly most powerful unbiased (UMPU). Taken together, these results suggest that the LTT-GLS, LM, and CLR tests are all asymptotically efficient under strong identification, subject to the stated regularity conditions.

Proposition S1 (Equivalence between LM and LTT-GLS under strong identification)**.** Suppose the process Z admits representation (2) and Assumptions 1′, 3 hold. Additionally, suppose there exists a sequence of stopping times $(T_m)_{m\geq 1}$ tending to infinity and a bounded sequence of constants $(K_m)_{m\geq 1}$ satisfying the following conditions for every $m \geq 1$: (i) for some constant $\mathfrak{s} \in (0,2)$, $\|b_t\| + \|\sigma_t\| + \|\sigma_t^{-1}\| + \int(\|x\|^{\mathfrak{s}} \wedge 1)F_t(\mathrm{d}x) \leq K_m$ for all $t \in [0, T_m]$, where $F_t(\cdot)$ is the spot Lévy measure of the jump process J; (ii) for some $\varkappa > 0$, $\mathbb{E}[|\sigma_{t\wedge T_m} - \sigma_{s\wedge T_m}|^2] \leq K_m|t-s|^{2\varkappa}$ for all $s,t < \infty$. Consider the local alternatives

$$\beta_n = \beta_0 + d\Delta_n^{1/2}, \qquad d \in \mathbb{R}. \tag{S5}$$

Recall $v_{S,p} = h_0^\intercal c_{\mathfrak{a}_p} h_0$ and $v_{T,p} = (g_0^\intercal c_{\mathfrak{a}_p}^{-1} g_0)^{-1}$ and let $J_p^X \equiv \Delta X_{\mathfrak{a}_p}$ be the p-th X-jump; define

$$\mathcal{I}^* \equiv \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{(J_p^X)^2}{v_{S,p}} > 0.$$

Let $\hat{c}_{n,i(p)} \equiv (\hat{c}_{n,i(p)-} + \hat{c}_{n,i(p)+})/2$ be consistent for $c_{\mathfrak{a}_p}$, and construct LM_n by replacing $(Z_p, c_{\mathfrak{a}_p})$ in (16)–(19) with $(\Delta_n^{-1/2}\Delta_{i(p)}^n Z, \hat{c}_{n,i(p)})$. Let the OLS estimator $\tilde{\beta}_n$ satisfy $\tilde{\beta}_n - \beta_n = o_{\mathbb{P}}(1)$ along each sequence in (S5), and set

$$\hat{v}_{S,n,p}^{\mathrm{GLS}} \equiv (-\tilde{\beta}_n, 1)\hat{c}_{n,i(p)}(-\tilde{\beta}_n, 1)^\intercal, \qquad \hat{\mathcal{I}}_n \equiv \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{(\Delta_{i(p)}^n X)^2}{\hat{v}_{S,n,p}^{\mathrm{GLS}}},$$

and define the squared LTT-GLS Wald statistic

$$W_{\text{LTT-GLS},n} \equiv \frac{\hat{\mathcal{I}}_n}{\Delta_n}(\hat{\beta}_n^{\mathrm{GLS}} - \beta_0)^2.$$

Then, for every fixed $d \in \mathbb{R}$, the following statements hold:

(a) The signed square roots of LM_n and $W_{\text{LTT-GLS},n}$ satisfy

$$K_{LM,n} = K_{\text{LTT-GLS},n} + o_{\mathbb{P}}(1) = \frac{1}{\sqrt{\mathcal{I}^*}} \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{J_p^X S_{n,p}(\beta_0)}{v_{S,p}} + o_{\mathbb{P}}(1), \tag{S6}$$

where $LM_n = K_{LM,n}^2$ and $W_{\text{LTT-GLS},n} = K_{\text{LTT-GLS},n}^2$. Consequently,

$$LM_n - W_{\text{LTT-GLS},n} = o_{\mathbb{P}}(1). \tag{S7}$$

(b) Conditionally on $\mathcal{F}$,

$$(LM_n, W_{\text{LTT-GLS},n}) \xrightarrow{\mathcal{L}\text{-}s} (Q_d, Q_d), \qquad Q_d \mid \mathcal{F} \sim \chi_1^2(d^2\mathcal{I}^*).$$

Hence, with $c_\alpha = \chi_{1,1-\alpha}^2$,

$$\mathbb{P}_{\beta_n}(1\{LM_n > c_\alpha\} \neq 1\{W_{\text{LTT-GLS},n} > c_\alpha\}) \longrightarrow 0.$$

(c) If, in addition, LTT's Assumptions 4 and 5 hold so that the LAMN argument

in Theorem 3 of Li, Todorov, and Tauchen (2017) applies to the selected event experiment (or its fixed-arrival conditional-Gaussian analogue applies), then the LTT semiparametric efficiency bound is

$$\Sigma^* = (\mathcal{I}^*)^{-1}.$$

Thus the two tests have the same local noncentrality parameter $d^2/\Sigma^* = d^2\mathcal{I}^*$ and are generated by the same semiparametrically efficient central sequence.

Proof of Proposition S1. An argument essentially identical to that of Lemma 1 in Li, Li, Patton, and Wei (2026), extended to the bivariate case, yields the following coupling result:

$$\Delta^n_{i(p)} Z = J_p^X g_n + \Delta_n^{1/2} \sigma_{\mathfrak{a}_p} \cdot \Psi_p + o_{\mathbb{P}}(\Delta_n^{1/2}), \quad \forall p \in \mathcal{P}_{\mathfrak{a}_B} \tag{S8}$$

where Ψ_p's are i.i.d. standard bivariate normals and $g_n = (1, \beta_n)^\intercal$. Put

$$\overline{Z}_{n,p} \equiv \Delta_n^{-1/2} \Delta^n_{i(p)} Z, \qquad S^0_{n,p} \equiv h_0^\intercal \overline{Z}_{n,p} = \frac{\Delta^n_{i(p)} Y - \beta_0 \Delta^n_{i(p)} X}{\sqrt{\Delta_n}} = S_{n,p}(\beta_0).$$

Since $h_0^\intercal g_0 = 0$, $h_0^\intercal e_2 = 1$, and $g_n = g_0 + d\sqrt{\Delta_n} e_2$, the coupling result in (S8) gives

$$S^0_{n,p} = d J_p^X + h_0^\intercal \sigma_{\mathfrak{a}_p} \Psi_p + o_{\mathbb{P}}(1). \tag{S9}$$

In particular, $S^0_{n,p} = O_{\mathbb{P}}(1)$ for every fixed d and for each p. First, rewrite the LM statistic. Denote $T^0_{n,p} \equiv g_0^\intercal \hat{c}^{-1}_{n,i(p)} \overline{Z}_{n,p} / (g_0^\intercal \hat{c}^{-1}_{n,i(p)} g_0), \hat{v}^0_{S,n,p} \equiv h_0^\intercal \hat{c}_{n,i(p)} h_0, \hat{v}^0_{T,n,p} \equiv (g_0^\intercal \hat{c}^{-1}_{n,i(p)} g_0)^{-1}, \hat{\lambda}^0_{n,p} \equiv \sqrt{\hat{v}^0_{T,n,p} / \hat{v}^0_{S,n,p}}$. Let $S^{0,*}_{n,p} = S^0_{n,p} / \sqrt{\hat{v}^0_{S,n,p}}$ and $T^{0,*}_{n,p} = T^0_{n,p} / \sqrt{\hat{v}^0_{T,n,p}}$. Then the numerator and denominator of the finite-sample version of the LM statistic in (19) satisfy the following identities:

$$\hat{\lambda}^0_{n,p} S^{0,*}_{n,p} T^{0,*}_{n,p} = \frac{T^0_{n,p} S^0_{n,p}}{\hat{v}^0_{S,n,p}}, \qquad (\hat{\lambda}^0_{n,p})^2 (T^{0,*}_{n,p})^2 = \frac{(T^0_{n,p})^2}{\hat{v}^0_{S,n,p}}$$

Define the restricted jump estimator on the original scale by $\hat{J}^{X,0}_{n,p} \equiv \sqrt{\Delta_n} T^0_{n,p}$. Using (S8),

$$\hat{J}^{X,0}_{n,p} = \frac{g_0^\intercal \hat{c}^{-1}_{n,i(p)} \Delta^n_{i(p)} Z}{g_0^\intercal \hat{c}^{-1}_{n,i(p)} g_0}$$

$$= J_p^X + \sqrt{\Delta_n}\frac{g_0^\intercal \hat{c}_{n,i(p)}^{-1}(dJ_p^X e_2 + \sigma_{\mathfrak{a}_p}\Psi_p + o_{\mathbb{P}}(1))}{g_0^\intercal \hat{c}_{n,i(p)}^{-1} g_0} = J_p^X + O_{\mathbb{P}}(\sqrt{\Delta_n}). \tag{S10}$$

Notice that volatility estimation error does not contaminate the leading jump term. Choosing the sign of the LM square root to agree with its score numerator gives

$$K_{LM,n} = \frac{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \hat{J}_{n,p}^{X,0} S_{n,p}^0 / \hat{v}_{S,n,p}^0}{\left\{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} (\hat{J}_{n,p}^{X,0})^2 / \hat{v}_{S,n,p}^0\right\}^{1/2}} = \frac{1}{\sqrt{\mathcal{I}^*}} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{J_p^X S_{n,p}^0}{v_{S,p}} + o_{\mathbb{P}}(1). \tag{S11}$$

This follows from (S10) and $\hat{v}_{S,n,p}^0 = v_{S,p} + o_{\mathbb{P}}(1)$.

Next, expand the LTT-GLS statistic. Because $\hat{c}_{n,p} \xrightarrow{\mathbb{P}} c_{\mathfrak{a}_p}$ and $\tilde{\beta}_n \xrightarrow{\mathbb{P}} \beta_0$ under (S5), LTT's optimal weight in their equation (4.5) satisfies

$$w^*(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}, \tilde{\beta}_n) = \frac{2}{(-\tilde{\beta}_n, 1)(\hat{c}_{n,i(p)-} + \hat{c}_{n,i(p)+})(-\tilde{\beta}_n, 1)^\intercal} = \frac{1}{\hat{v}_{S,n,p}^{\mathrm{GLS}}} \xrightarrow{\mathbb{P}} \frac{1}{v_{S,p}} \tag{S12}$$

Moreover, (S8) gives $\Delta_{i(p)}^n X = J_p^X + O_{\mathbb{P}}(\sqrt{\Delta_n})$ and hence $\hat{\mathcal{I}}_n \xrightarrow{\mathbb{P}} \mathcal{I}^*$. The WLS identity in (9) now gives

$$\Delta_n^{-1/2}(\hat{\beta}_n(\mathfrak{a}_B, w^*) - \beta_0) = \frac{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} w^*(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}, \tilde{\beta}_n)\Delta_{i(p)}^n X \cdot S_{n,p}^0}{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} w^*(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}, \tilde{\beta}_n)\Delta_{i(p)}^n X^2}$$

So the above and (S12) imply

$$\begin{aligned} K_{\text{LTT-GLS},n} &= \sqrt{\frac{\hat{\mathcal{I}}_n}{\Delta_n}}\left(\hat{\beta}_n(\mathfrak{a}_B, w^*) - \beta_0\right) \\ &= \frac{\sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \Delta_{i(p)}^n X \cdot S_{n,p} / \hat{v}_{S,n,p}^{\mathrm{GLS}}}{\sqrt{\hat{\mathcal{I}}_n}} = \frac{1}{\sqrt{\mathcal{I}^*}} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{J_p^X \cdot S_{n,p}^0}{v_{S,p}} + o_{\mathbb{P}}(1). \end{aligned} \tag{S13}$$

Equations (S11) and (S13) prove (S6). Their common leading term is $O_{\mathbb{P}}(1)$ by (S9), so squaring proves (S7).

Conditionally on $\mathcal{F}$, (S9) implies

$$\frac{1}{\sqrt{\mathcal{I}^*}} \sum_{p\in\mathcal{P}_{\mathfrak{a}_B}} \frac{J_p^X \cdot S_{n,p}^0}{v_{S,p}} \xrightarrow{\mathcal{L}\text{-}s} \mathcal{N}(d\sqrt{\mathcal{I}^*}, 1),$$

because the conditional variance of $\sum_p J_p^X h_0^\intercal \sigma_{\mathfrak{a}_p} \Psi_p$ is $\sum_p (J_p^X)^2 / v_{S,p} = \mathcal{I}^*$. Part (b) follows by squaring. Since the limiting distribution is continuous at c_α, the difference between the two rejection indicators converges to zero in probability.

Finally, LTT's optimal asymptotic variance in their equation (4.8) is

$$\Sigma^* = \left\{ \sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{2(J_p^X)^2}{h_0^\intercal (c_{\mathfrak{a}_p -} + c_{\mathfrak{a}_p}) h_0} \right\}^{-1} = \left(\sum_{p \in \mathcal{P}_{\mathfrak{a}_B}} \frac{(J_p^X)^2}{v_{S,p}} \right)^{-1} = (\mathcal{I}^*)^{-1}.$$

Under the LAMN restrictions of LTT's semiparametric efficiency analysis, their Theorem 3 shows that this bound is sharp. Equivalently, profiling the unrestricted nuisance jump size J_p from the pth conditional Gaussian likelihood gives the Schur-complement information

$$(J_p^X)^2 \left\{ e_2^\intercal c_p^{-1} e_2 - \frac{(e_2^\intercal c_p^{-1} g_0)^2}{g_0^\intercal c_p^{-1} g_0} \right\} = \frac{(J_p^X)^2}{h_0^\intercal c_p h_0} = \frac{(J_p^X)^2}{v_{S,p}}.$$

Summing over announcements yields $\mathcal{I}^*$, which proves part (c). □

Proposition S2. In the Gaussian limit experiment (15) with $P = |\mathcal{P}_{\mathfrak{a}_B}|$, let $\lambda_p > 0$, θ_p, $\xi_p^* \equiv \xi_p / \sqrt{v_{T,p}}$ and $\delta = \beta - \beta_0$ be as in Theorem 4, $\Lambda = \mathrm{diag}((\lambda_p)_p)$, and $\pi(\tau) \equiv \mathbb{P}\{\chi_1^2(\tau) > \chi_{1,1-\alpha}^2\}$. Consider drifting parameters under strong or semi-strong identification and the local alternative

$$\xi_p^* = r_n u_p, \qquad r_n \to \infty, \qquad \delta_n = d / r_n, \qquad d \in \mathbb{R}, \tag{S14}$$

for fixed $u = (u_1, \ldots, u_P)^\intercal \in \mathbb{R}_*^P$, and define

$$w_p \equiv \frac{u_p^2}{\sum_q u_q^2}, \qquad \mathcal{R}(w, \lambda) \equiv \frac{\left(\sum_p w_p \lambda_p \right)^2}{\sum_p w_p \lambda_p^2}, \qquad \tau \equiv d^2 \sum_p \lambda_p^2 u_p^2.$$

(a) Along the local alternative (S14), the tests φ_{LM} and φ_{CLR} with critical values computed conditionally on T^* are asymptotically UMPU, and the local power functions satisfy $\lim_n \mathbb{E}_{\delta_n, \xi^*}[\varphi_{LM}] = \lim_n \mathbb{E}_{\delta_n, \xi^*}[\varphi_{CLR}] = \pi(\tau)$.

(b) The statistics $LM^h = Q_{ST}^2 / Q_T$ and LR^h of Theorem 4(b), each with its exact conditional critical value, satisfy $\lim_n \mathbb{E}_{\delta_n, \xi^*}[\varphi^h] = \pi\big(\mathcal{R}(w, \lambda) \cdot \tau\big)$.

(c) With $\lambda_{\min} \equiv \min_p \lambda_p$, $\lambda_{\max} \equiv \max_p \lambda_p$ (the extrema may be restricted to $\{p : w_p > 0\}$),

$$1 \;\geq\; \mathcal{R}(w,\lambda) \;\geq\; \mathcal{R}_\lambda \equiv \frac{4\lambda_{\min}\lambda_{\max}}{(\lambda_{\min}+\lambda_{\max})^2}.$$

If $\lambda_{\min} < \lambda_{\max}$, equality on the left holds iff λ_p is constant on $\{p : w_p > 0\}$, and equality on the right holds iff w puts mass $\lambda_{\max}/(\lambda_{\min}+\lambda_{\max})$ on $\{p : \lambda_p = \lambda_{\min}\}$ and the remainder on $\{p : \lambda_p = \lambda_{\max}\}$. If $\lambda_{\min} = \lambda_{\max}$, both inequalities are equalities. Thus $\mathcal{R}_\lambda$ is sharp uniformly over the unknown jump-size configuration.

(d) Since π is strictly increasing,

$$\begin{aligned} 0 \;\leq\; \lim_n \mathbb{E}[\varphi_{CLR}] - \lim_n \mathbb{E}[\varphi_{CLR^h}] &= \pi(\tau) - \pi\big(\mathcal{R}(w,\lambda)\tau\big) \\ &\leq\; G(\mathcal{R}_\lambda) \;\leq\; C_\alpha \log(1/\mathcal{R}_\lambda), \end{aligned}$$

where $G : (0,1] \to [0, 1-\alpha)$ is defined by $G(\kappa) \equiv \sup_{t\geq 0}\{\pi(t) - \pi(\kappa t)\}$ and is strictly decreasing, with $G(1) = 0$, and $C_\alpha \equiv \sup_{t\geq 0} t\,\pi'(t)$; at $\alpha = 0.05$, $C_\alpha \approx 0.435$ and, e.g., $G(0.99) \approx 0.004$, $G(0.9) \approx 0.046$, $G(0.75) \approx 0.124$, $G(0.5) \approx 0.292$, $G(0.19) \approx 0.609$.

Proof of Proposition S2. (a) Write the generic term of (20) as $f_p(\delta') = \{\delta'\lambda_p S_p^* + (1+\delta'\theta_p)T_p^*\}^2/\{\delta'^2\lambda_p^2 + (1+\delta'\theta_p)^2\}$, so $LR = \sup_{\delta'\in\overline{\mathbb{R}}} \sum_p f_p(\delta') - Q_T$. Two elementary facts drive the argument. First, for every δ' (including $\delta' = \pm\infty$), put $a = \delta'\lambda_p$, $b = 1+\delta'\theta_p$, and $D = a^2+b^2$. Then $g_p(\delta') = a^2/D$, $b^2/D = 1 - g_p(\delta')$, and $|ab|/D = \{g_p(\delta')(1-g_p(\delta'))\}^{1/2}$. Expanding the square gives

$$\begin{aligned} f_p(\delta') - (T_p^*)^2 = \frac{(aS_p^* + bT_p^*)^2 - (a^2+b^2)(T_p^*)^2}{D} &= \frac{a^2[(S_p^*)^2 - (T_p^*)^2] + 2abS_p^*T_p^*}{D} \\ &= g_p(\delta')\{(S_p^*)^2 - (T_p^*)^2\} + 2\frac{ab}{D}S_p^*T_p^*. \end{aligned}$$

The desired bound is just a control of the cross term: indeed,

$$\begin{aligned} &2(S_p^*)^2 - \tfrac{1}{2}g_p(\delta')(T_p^*)^2 - \{f_p(\delta') - (T_p^*)^2\} \\ &\quad = \{2 - g_p(\delta')\}(S_p^*)^2 + \tfrac{1}{2}g_p(\delta')(T_p^*)^2 - 2\frac{ab}{D}S_p^*T_p^* \;\geq\; 0, \end{aligned}$$

because the discriminant of the quadratic polynomial is non-positive. Hence

$$f_p(\delta') - (T_p^*)^2 \;\le\; 2(S_p^*)^2 - \tfrac{1}{2}\, g_p(\delta')(T_p^*)^2, \qquad g_p(\delta') \equiv \frac{\delta'^2\lambda_p^2}{\delta'^2\lambda_p^2 + (1+\delta'\theta_p)^2}, \tag{S15}$$

where g_p is continuous on $\overline{\mathbb{R}}$ and vanishes only at $\delta' = 0$. To see why this concentrates the profiled maximizer near zero, sum (S15) over p:

$$\sum_p f_p(\delta') - Q_T \le 2\sum_p (S_p^*)^2 - \frac{1}{2}\sum_p g_p(\delta')(T_p^*)^2.$$

The scale λ_p in the first display below comes only from rewriting the mean of S_p^* in terms of the standardized jump size:

$$\mathbb{E}_{\delta,\xi^*}[S_p^*] = \frac{\delta\xi_p}{\sqrt{v_{S,p}}} = \delta\sqrt{\frac{v_{T,p}}{v_{S,p}}}\xi_p^* = \delta\lambda_p\xi_p^*, \qquad \mathbb{E}_{\delta,\xi^*}[T_p^*] = (1+\delta\theta_p)\xi_p^*.$$

Thus, along (S14),

$$S_p^* = d\lambda_p u_p + O_{\mathbb{P}}(1), \qquad T_p^* = r_n u_p + d\theta_p u_p + O_{\mathbb{P}}(1) = r_n u_p + O_{\mathbb{P}}(1),$$

so $\sum_p (S_p^*)^2 = O_{\mathbb{P}}(1)$. The negative term, by contrast, is large once $|\delta'|$ is larger than the local scale $1/r_n$. Indeed, on the local annulus $M/r_n \le |\delta'| \le \varepsilon$, for ε small enough, $g_p(\delta') \ge c\,\delta'^2\lambda_p^2$ uniformly in p for some $c > 0$, and therefore

$$\sum_p g_p(\delta')(T_p^*)^2 \ge c\, r_n^2\delta'^2 \sum_p \lambda_p^2 u_p^2 + o_{\mathbb{P}}(r_n^2\delta'^2) \ge cM^2 \sum_p \lambda_p^2 u_p^2 + o_{\mathbb{P}}(1).$$

Outside any fixed neighborhood of zero, continuity of g_p and $u \neq 0$ make $\sum_p g_p(\delta')u_p^2$ bounded away from zero, so the same negative term is of order r_n^2. Hence, after choosing M large, with probability approaching one,

$$\sup_{|\delta'|\ge M/r_n} \left\{ \sum_p f_p(\delta') - Q_T \right\} < 0.$$

Since the objective equals zero at $\delta' = 0$, the profiled supremum is attained, up to an

$o_{\mathbb{P}}(1)$ error, inside $|\delta'| \le M/r_n$. Second, use the exact identity

$$f_p(\delta') - (T_p^*)^2 = \frac{\delta'\lambda_p\big\{2(1+\delta'\theta_p)S_p^*T_p^* + \delta'\lambda_p\big((S_p^*)^2 - (T_p^*)^2\big)\big\}}{\delta'^2\lambda_p^2 + (1+\delta'\theta_p)^2}.$$

Evaluating at $\delta' = h/r_n$, uniformly over $|h| \le M$, the denominator $\delta'^2\lambda_p^2 + (1+\delta'\theta_p)^2 = 1 + 2h\theta_p/r_n + h^2(\theta_p^2+\lambda_p^2)/r_n^2 = 1 + o(1)$ and

$$\begin{aligned}
\frac{h\lambda_p}{r_n}\,2\Big(1 + \frac{h\theta_p}{r_n}\Big)S_p^*T_p^* &= 2h\lambda_p u_p S_p^* + o_{\mathbb{P}}(1),\\
\frac{h^2\lambda_p^2}{r_n^2}\big((S_p^*)^2 - (T_p^*)^2\big) &= -h^2\lambda_p^2 u_p^2 + o_{\mathbb{P}}(1),
\end{aligned}$$

because $T_p^*/r_n \xrightarrow{\mathbb{P}} u_p$ and $(S_p^*)^2/r_n^2 \xrightarrow{\mathbb{P}} 0$. Summing over p yields

$$\sum_p f_p(h/r_n) - Q_T = 2h\langle \Lambda u, S^*\rangle - h^2\|\Lambda u\|^2 + o_{\mathbb{P}}(1).$$

Maximizing the limiting quadratic in h gives

$$LR = \frac{\langle \Lambda u, S^*\rangle^2}{\|\Lambda u\|^2} + o_{\mathbb{P}}(1).$$

And note

$$LM = \frac{\big(\sum_p \lambda_p S_p^*T_p^*\big)^2}{\sum_p \lambda_p^2 (T_p^*)^2} = \frac{\big(r_n\sum_p \lambda_p u_p S_p^* + o_{\mathbb{P}}(r_n)\big)^2}{r_n^2\sum_p \lambda_p^2 u_p^2 + o_{\mathbb{P}}(r_n^2)} = \frac{\langle \Lambda u, S^*\rangle^2}{\|\Lambda u\|^2} + o_{\mathbb{P}}(1).$$

Thus $LR = LM + o_{\mathbb{P}}(1)$. Since $S^* \sim \mathcal{N}(d\Lambda u, I)$ along (S14), $LM \xrightarrow{\mathcal{L}} \chi^2_1(\tau)$ with noncentrality parameter $\tau = d^2\|\Lambda u\|^2$. The same expansion under H_0 conditionally on $T^* = t_n^*$ with $t_n^*/r_n \to u$ yields $LR \mid T^* \xrightarrow{\mathcal{L}} \chi^2_1$ and hence $cv^{\alpha}_{CLR}(T^*) \xrightarrow{\mathbb{P}} \chi^2_{1,1-\alpha}$; the LM test needs no argument since $LM \mid T^* \sim \chi^2_1$ exactly. It remains to spell out the efficiency claim. Along (S14), conditioning on T^* asymptotically fixes the nuisance direction, because $T^*/r_n \xrightarrow{\mathbb{P}} u$. The relevant limit experiment for the local parameter d is therefore the Gaussian shift

$$S^* = d\,\Lambda u + \varepsilon, \qquad \varepsilon \sim \mathcal{N}(0, I_P).$$

All information about d is contained in the scalar projection

$$Z_u \equiv \frac{\langle \Lambda u, S^* \rangle}{\|\Lambda u\|}, \qquad Z_u \sim \mathcal{N}(d\|\Lambda u\|, 1),$$

since the component of S^* orthogonal to Λu has a distribution that does not depend on d. Equivalently, the log-likelihood ratio in the limit experiment is

$$\log \frac{\mathrm{d}\mathbb{P}_d}{\mathrm{d}\mathbb{P}_0}(S^*) = d\langle \Lambda u, S^* \rangle - \frac{1}{2} d^2 \|\Lambda u\|^2.$$

Thus the efficient information is $\|\Lambda u\|^2$ and the local noncentrality is $d^2\|\Lambda u\|^2 = \tau = \lim_n \delta_n^2 \sum_p \xi_p^2 / v_{S,p}$. For the two-sided normal mean problem of testing $H_0 : \eta = 0$ against $H_1 : \eta \neq 0$ given $Z_u \sim \mathcal{N}(\eta, 1)$, the following identities verify the two conditions of Theorem 4.4.1 of Lehmann and Romano (2005):

$$\begin{aligned} &\mathbb{P}_0(Z_u < z_{\alpha/2} \text{ or } Z_u > z_{1-\alpha/2}) = \alpha \\ &\mathbb{E}_{H_0}[Z_u 1\{Z_u < z_{\alpha/2} \text{ or } Z_u > z_{1-\alpha/2}\}] = 0 \end{aligned}$$

where $z_{\alpha/2}$ and $z_{1-\alpha/2}$ are the lower and upper $\alpha/2$-quantiles of the standard normal distribution, respectively. Therefore, the equivalent chi-square test $1\{Z_u^2 > \chi^2_{1,1-\alpha}\}$ is the uniformly most powerful unbiased (UMPU) level-α test with the local power function $\mathbb{P}\{\chi_1^2(\tau) > \chi^2_{1,1-\alpha}\} = \pi(\tau)$.

(b) In the semi-strong/strong identification cases, $Q_T \to \infty$. Moreover, $LR^h = LM^h + o_{\mathbb{P}}(1)$ and $cv^{\alpha}_{CLR^h}(Q_T) \to \chi^2_{1,1-\alpha}$ follow from (22).

(c) With $\lambda_{\min} \equiv \min_p \lambda_p$ and $\lambda_{\max} \equiv \max_p \lambda_p$, it follows that

$$\underline{\mathcal{R}}_\lambda \equiv \frac{4\lambda_{\min}\lambda_{\max}}{(\lambda_{\min} + \lambda_{\max})^2} \overset{(i)}{\leq} \mathcal{R}(w, \lambda) = \frac{(\sum_p w_p \lambda_p)^2}{\sum_p w_p \lambda_p^2} \overset{(ii)}{\leq} 1.$$

To prove inequality (i), let $\bar{\lambda} \equiv \sum_p w_p \lambda_p$. Since $\lambda_{\min} > 0$, $\sum_p w_p = 1$, and $\lambda_{\min} \leq \lambda_p \leq \lambda_{\max}$ on the support of w, it follows that

$$0 \leq (\lambda_p - \lambda_{\min})(\lambda_{\max} - \lambda_p) \quad \Longrightarrow \quad \lambda_p^2 \leq (\lambda_{\min} + \lambda_{\max})\lambda_p - \lambda_{\min}\lambda_{\max}.$$

Multiplying by w_p and summing over p therefore yields

$$\sum_p w_p \lambda_p^2 \le (\lambda_{\min} + \lambda_{\max})\bar{\lambda} - \lambda_{\min}\lambda_{\max}.$$

Consequently,

$$\begin{aligned}\mathcal{R}(w,\lambda) &\ge \frac{\bar{\lambda}^2}{(\lambda_{\min}+\lambda_{\max})\bar{\lambda} - \lambda_{\min}\lambda_{\max}}, \\ &\frac{\bar{\lambda}^2}{(\lambda_{\min}+\lambda_{\max})\bar{\lambda} - \lambda_{\min}\lambda_{\max}} - \frac{4\lambda_{\min}\lambda_{\max}}{(\lambda_{\min}+\lambda_{\max})^2} \\ &= \frac{\big((\lambda_{\min}+\lambda_{\max})\bar{\lambda} - 2\lambda_{\min}\lambda_{\max}\big)^2}{(\lambda_{\min}+\lambda_{\max})^2\big((\lambda_{\min}+\lambda_{\max})\bar{\lambda} - \lambda_{\min}\lambda_{\max}\big)} \ge 0.\end{aligned}$$

The denominators are positive because $\lambda_{\min} > 0$ and $\bar{\lambda} \in [\lambda_{\min}, \lambda_{\max}]$. This proves inequality (i). If $\lambda_{\min} < \lambda_{\max}$, equality holds iff $\lambda_p \in \{\lambda_{\min}, \lambda_{\max}\}$ on the support of w and

$$\bar{\lambda} = \frac{2\lambda_{\min}\lambda_{\max}}{\lambda_{\min}+\lambda_{\max}},$$

or, equivalently, the total weights assigned to $\lambda_{\min}$ and $\lambda_{\max}$ are $\lambda_{\max}/(\lambda_{\min}+\lambda_{\max})$ and $\lambda_{\min}/(\lambda_{\min}+\lambda_{\max})$, respectively. If $\lambda_{\min} = \lambda_{\max}$, the bound holds with equality trivially. Finally, weighted Cauchy–Schwarz gives

$$\left(\sum_p w_p\lambda_p\right)^2 \le \left(\sum_p w_p\right)\left(\sum_p w_p\lambda_p^2\right) = \sum_p w_p\lambda_p^2,$$

which proves $\mathcal{R}(w,\lambda) \le 1$. Equality in Cauchy–Schwarz holds iff λ_p is constant on the support of w.

(d) π is strictly increasing because $\chi_1^2(\tau)$ is stochastically increasing in τ. For the log bound, $\pi(t) - \pi(\kappa t) = \int_\kappa^1 t\,\pi'(ts)\,ds \le C_\alpha \log(1/\kappa)$, using $\pi'(\tau) = \frac{1}{2}\{\mathbb{P}(\chi_3^2(\tau) > \chi_{1,1-\alpha}^2) - \mathbb{P}(\chi_1^2(\tau) > \chi_{1,1-\alpha}^2)\}$. □

S4. More on Empirical Applications

S4.1 Confidence Sets for the Fed Information Effect

Table S1: Identification-robust 99% confidence sets for policy-news responses

FOMC day	A. Real yield	B. Dividend price	C. Stock return	Method
2025-09-17	$[0.429, 0.937]$	$[-2.189, 13.218]$	$[-0.077, 0.018]$	AR
2025-10-29	$[0.436, 1.222]$	$[-16.086, 0.070]$	$[-0.003, 0.093]$	AR
2025-12-10	$[0.362, 2.606]$	$[-9.583, 15.780]$	$[-0.528, -0.106]$	AR
2026-01-28	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	AR
2026-03-18	$[0.117, 0.967]$	$[-7.164, 4.227]$	$[-0.113, 0.023]$	AR
2026-04-29	$\mathbb{R}$	$\mathbb{R} \setminus (-2.359, 12.570)$	$\mathbb{R} \setminus (-0.020, 0.051)$	AR
2026-06-17	$[0.584, 0.875]$	$[-0.881, 5.283]$	$[-0.084, -0.031]$	AR
2026-07-29	$[0.005, 0.226]$	$[1.139, 5.421]$	$[-0.050, -0.027]$	AR
2026-09-16	$[-0.651, 0.174]$	$[-1.810, 5.798]$	$[-0.052, 0.010]$	AR
Pooled	$[0.435, 0.572]$	$[0.61, 3.36]$	$[-0.049, -0.029]$	CLR

Notes: Announcement-specific sets use AR with χ^2_1 critical values; pooled sets use CLR with a common β across 9 meetings. All panels use X with the normalization described in Figure 5. Units are bp (panel A) or log-return percentage points (panels B and C) per SOFR5-equivalent bp. Panel B uses 180-day dividend prices inferred from put–call parity; Panel C uses the S&P 500 index. Covariance estimates use $\kappa = 0$ and $k_n = 30$ (panel A) or 28 (panels B and C); PCA weights are treated as fixed. Table 2 reports the corresponding 95% sets.

S4.2 War-Related Oil News

Tables S2 and S3 report the news events in panels (a) and (b) of Figure 6, respectively.

Table S2: Oil news related to the Russia–Ukraine war

				Return (%)	
	Date	ET	News and source	WTI	TLT
•	2022-02-11	13:32†	U.S. believes Putin has decided to invade (pre-invasion) PBS reporter on Twitter; news at 13:28	0.73	0.073
○	2022-03-03	12:40	Round-two talks end: only humanitarian corridors agreed Podolyak on Twitter	−0.10	0.008
○	2022-03-07	12:50	Round-three talks end: small progress on corridors Podolyak on Twitter	−0.17	−0.039
○	2022-03-14	09:55	Talks pause until the next day; working groups go on Podolyak on Twitter	−0.35	−0.075
○	2022-03-15	15:36	Day's talks end: deep differences, room for compromise Podolyak on Twitter	0.02	−0.008
•	2022-03-16	10:15†	Progress on a 15-point ceasefire and neutrality plan FT report, relayed by LiveSquawk; news at 10:12	−0.77	−0.067
○	2022-03-16	12:45	Kyiv: the FT draft reflects only Russian demands Podolyak on Twitter	0.07	0.008
○	2022-11-15	13:41	Russian missiles crossed into Poland, killing two people AP alert citing a U.S. official; later corrected	−0.08	0.011
○	2023-07-19	11:10	Ukraine-bound ships deemed potential military carriers Interfax (Russian Defense Ministry)	−0.05	−0.031
○	2025-10-16	13:11	Trump announces a summit with Putin in Budapest Trump on Truth Social	−0.12	−0.016
○	2025-12-15	11:03	U.S. official: Berlin talks on Ukraine very positive Reuters alert	−0.04	0.030

Notes: Events plotted in Figure 6, panel (a). Times are New York time and denote the start of the one-minute return interval. Returns are in percent, rounded to two decimal places for WTI and three for TLT. Filled and open circles denote WTI returns above and below truncation, respectively. † marks the first WTI threshold crossing after the news; the original news times are given in the source lines.

Table S3: Oil news related to the Iran war

	Date	ET	News and source	Return (%) WTI	Return (%) TLT
○	2026-03-11	10:05	IEA members agree to release 400 million barrels Reuters alert (IEA)	−0.29	0.016
○	2026-03-18	09:59	Iranian gas flows to Iraq halted after South Pars attack Reuters alert (Iraqi official)	0.12	0.025
○	2026-03-18	14:35	QatarEnergy: extensive damage from Ras Laffan attack Reuters alert	−0.04	0.003
○	2026-04-08	10:31	Fars: Hormuz tankers halted over alleged Israeli breach Reuters alert (Fars news agency)	0.34	−0.015
○	2026-06-12	12:14	Pakistan PM: final text of Iran-U.S. peace deal agreed Reuters alert	0.14	0.012
○	2026-06-15	11:36	Senior U.S. official: U.S. has signed an MOU with Iran Reuters alert	0.00	−0.007
○	2026-07-07	14:46	U.S. revokes Iran-related General License X Reuters alert (Treasury website)	0.20	−0.002
•	2026-07-13	10:16	Trump reinstates the blockade of Iran Trump on Truth Social	1.14	−0.178
•	2026-08-25	11:33	Oman, Iran discuss phased reopening of Hormuz navigation Reuters alert (joint statement)	−0.35	0.044
○	2026-10-01	12:36	New U.S. sanctions on Iran, aimed at rail and auto groups Reuters alert (Treasury website)	0.18	−0.032

Notes: Events plotted in Figure 6, panel (b). Times are New York time and denote the start of the one-minute return interval. Returns are in percent, rounded to two decimal places for WTI and three for TLT. Filled and open circles denote WTI returns above and below truncation, respectively.

S5. Additional Simulation Results

Table S4 repeats Table 1 with oracle spot covariances on the same simulated paths. AR and conditional-test rejection rates fall from 5.42–6.90% to 4.67–4.99%, close to the 5% nominal level, supporting the attribution of their mild size distortion primarily to finite-sample error in large-k covariance estimation. LTT remains substantially oversized under weak identification.

Table S4: Null rejection percentages at the 5% nominal level with oracle spot volatilities.

r	AR	LM/K	CLR	CLR^h	CLC-$\hat{\mathfrak{c}}_{\mathrm{MLE}}$	CLC-$\hat{\mathfrak{c}}_{\mathrm{PP}}$	CLC-$\hat{\mathfrak{c}}_{\mathrm{KRS}}$	LTT–GLS	LTT–OLS
0	4.68	4.91	4.87	4.71	4.99	4.99	4.99	5.62	5.12
0.1	4.68	4.84	4.79	4.72	4.92	4.92	4.92	6.09	5.90
0.3	4.68	4.94	4.78	4.72	4.95	4.95	4.95	11.41	10.46
0.5	4.68	4.73	4.73	4.67	4.74	4.74	4.71	30.54	25.92

Notes: Entries repeat Table 1 on the same 10,000 paths with the true spot covariances $c_{\mathfrak{a}_p}$ and $c_{\mathfrak{a}_p-}$ of the discretized benchmark DGP in place of $\hat{c}_{n,i(p)+}$ and $\hat{c}_{n,i(p)-}$; critical values, simulation draws, and random seeds are unchanged. Monte Carlo standard errors are about 0.22 percentage points for entries near 5%.

References

Aliprantis, C. D., and K. C. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*. Springer, Berlin, 3 edn.

Birnbaum, A. (1955): "Characterizations of Complete Classes of Tests of Some Multiparametric Hypotheses, with Applications to Likelihood Ratio Tests," *The Annals of Mathematical Statistics*, 26(1), 21–36.

Lehmann, E. L., and J. P. Romano (2005): *Testing Statistical Hypotheses*. Springer, 3rd edn.

Li, J., Q. Li, A. J. Patton, and J. Wei (2026): "A Fixed-k Bipower Jump Test," Working paper.

Li, J., V. Todorov, and G. Tauchen (2017): "Jump Regressions," *Econometrica*, 85(1), 173–195.

Shao, J. (2003): *Mathematical Statistics*. Springer, 2 edn.