Unit Root Test with High-Frequency Data*

Sébastien Laurent † and Shuping Shi‡

†Aix-Marseille University
‡Macquarie University

May 10, 2020

Abstract

Deviations of asset prices from the random walk dynamic imply the predictability of asset returns and thus have important implications for portfolio construction and risk management. This paper proposes a real-time monitoring device for such deviations using intraday high-frequency data. The proposed procedures are based on unit root tests with in-fill asymptotics but extended to take the empirical features of high-frequency financial data (particularly jumps) into consideration. We derive the limiting distributions of the tests under both the null hypothesis of a random walk with jumps and the alternative of mean reversion/explosiveness with jumps. The limiting results show that ignoring the presence of jumps could potentially lead to severe size distortions of both the standard left-sided (against mean reversion) and right-sided (against explosiveness) unit root tests. The simulation results reveal satisfactory performance of the tests even with data from a relatively short time span. As an illustration, we apply the procedure to the Nasdaq composite index at the 10-minute frequency over two periods: around the peak of the dot-com bubble and during the 2015-2106 stock market sell-off. We find strong evidence of explosiveness dynamics in asset prices in late 1999 and mean reversion in late 2015. We also show that accounting for jumps when testing the random walk hypothesis on intraday data is empirically relevant and that ignoring jumps can lead to different conclusions.

Keywords: Unit root test, random walk, in-fill asymptotic, jumps, GARCH, periodicity, microstructure noise
JEL classification: C12, C22.

1 Introduction

The issue of whether stock prices follow a random walk or a mean-reverting process received considerable attention at the end of the 20th century. Evidence of mean reversion in stock prices or autocorrelation in long-horizon returns have been documented in the stock prices of the US (Fama

*The authors gratefully acknowledge Jun Yu, Olivier Scaillet, Xiaohu Wang, and participants at the QFFE 2019 conference in Marseille for helpful discussions. We thank the coeditor Eric Renault and three anonymous referees for very useful comments. Shi acknowledges research support from the Australian Research Council under project No. DE190100840. Laurent acknowledges research support from the French National Research Agency Grant ANR-17-EURE-0020. Sébastien Laurent, Aix-Marseille University (Aix-Marseille School of Economics), CNRS & EHESS, Aix-Marseille Graduate School of Management - IAE, France; E-mail: sebastien.laurent@univ-amu.fr. Shuping Shi, Department of Economics, Macquarie University; E-mail: shuping.shi@mq.edu.au.

and French, 1988; Poterba and Summers, 1988; Lo and MacKinlay, 1988)[1] and many other countries (Richards, 1997; Balvers et al., 2000; Chaudhuri and Wu, 2003). There is also a burgeoning research program searching for evidence of asset prices deviating to an explosive regime (viz. speculative bubbles). It is argued that asset prices are explosive (Diba and Grossman, 1988) in the presence of speculative bubbles, as opposed to being a random walk under normal market conditions. With recently developed bubble identification techniques, the literature presents abundant evidence of explosiveness in asset prices.[2] It is important to note that the empirical evidence of both mean reversion and explosiveness of asset prices is observed from low-frequency (weekly, monthly or quarterly) data.

Several trading strategies have been developed to exploit the mean-reverting behavior (Balvers et al., 2000; Gatev et al., 2006; Serban, 2010) and the explosive dynamics (Brooks and Katsaris, 2005; Guenster and Kole, 2013; Milunovich et al., 2017) of asset prices. These trading strategies are shown to outperform the buy-and-hold strategy, with or without the consideration of transaction costs. They are, however, designed for low-frequency trading, which often requires a long holding period to be profitable. The readily available high-frequency financial data provide a strong motive for investors to extend those strategies to high-frequency settings and trade more frequently. The profitability of such high-frequency trading will rely critically upon having a timely and accurate identification technique for such deviations.

Moreover, deviations from the random walk imply the presence of a nonzero drift in a linear drift diffusion (e.g., Ornstein-Uhlenbeck) process. Laurent and Shi (2020) show that the presence of a nonzero drift results in the overestimation of the integrated variance using various realized volatility estimators (including jump robust estimators) and a power loss for jump detection procedures. As a remedy, they suggest using centered returns for the calculation of integrated volatilities and the construction of the jump test statistics. An effective tool for identifying such deviations in the high-frequency regime will, therefore, be an essential step for statistically documenting empirical evidence of nonzero drifts in the price dynamics of various assets and hence justifying the need for handling drifts with care.

This paper addresses this need by providing a real-time monitoring technique for deviations of asset prices from the random walk using high-frequency data. The real-time monitoring procedure arises from the unit root testing literature,[3] which started in the late 1970s and was catalyzed by the work of Nelson and Plosser (1982). The view that most economic time series are characterized by stochastic trends has since become prevalent. Despite the popularity of unit root testing, there is a profound concern regarding structural breaks caused by changes in institutional or policy settings (Kim et al., 1991). At the turn of the 20^{th} century, an enormous amount of effort was devoted to tackling this issue, considering different break types (such as breaks in the null or in the alternative, breaks in the mean, trend, or slope, and sudden or gradual breaks), known or unknown break dates, and the number of breaks.[4] Although one could employ a procedure to endogenously determine the break dates, some assumptions on the nature of the break (e.g., break numbers or break in mean, slope or trend) must be made for practical implementation. Those assumptions are often critical

[1]This finding is, however, subject to criticisms. See Lo and MacKinlay (1988); Richardson (1993); McQueen (1992); Kim et al. (1991); Miller et al. (1994).

[2]See, for example, Brooks and Katsaris (2005); Phillips et al. (2011); Phillips and Yu (2011); Homm and Breitung (2012); Phillips and Yu (2013); Phillips et al. (2015a); Milunovich et al. (2017); Narayan et al. (2016); Shi and Song (2016); Harvey et al. (2019).

[3]See, for example, Dickey and Fuller (1979, 1981); Said and Dickey (1984); Phillips (1987a); Phillips and Perron (1988); Kwiatkowski et al. (1992); Schmidt and Phillips (1992).

[4]See Perron (1989, 1990); Banerjee et al. (1992); Perron (1997); Lumsdaine and Papell (1997); Vogelsang and Perron (1998); Clemente et al. (1998); Lee and Strazicich (2001); Zivot and Andrews (2002); Lee and Strazicich (2003), among others.

and could lead to distinct results from those obtained under other choices. The unsatisfactory performance of those tests, therefore, prevents their widespread application.

The technique proposed here utilizes intraday data from a relatively short time interval, which is in sharp contrast to the existing literature that searches for evidence of deviations with low-frequency data and usually over a long time span. Therefore, unlike conventional unit root tests, structural breaks are of less concern for the new test. Additionally, the use of intraday data could potentially enable more effective detection of such deviations. The unit root test for high-frequency data employs in-fill asymptotics, where the sample period N is fixed, and the sampling interval Δ converges to zero. The analysis of a fixed time span and fine sampling intervals is typical in the high-frequency literature (e.g., Merton, 1980; Andersen and Bollerslev, 1998a). Moreover, in-fill asymptotics have been shown to provide better approximations to their finite sample counterparts (Yu, 2014; Zhou and Yu, 2015; Jiang et al., 2017, 2018) than long-span ($N \to \infty$) and double asymptotics ($N \to \infty$ and $\Delta \to 0$).

Although the in-fill asymptotic of unit root tests was developed as early as 1987 (Phillips, 1987a; Perron, 1991), there have been very few attempts at applying the test to high-frequency data over the past three decades. This is partially due to the paper by Shiller and Perron (1985), who show through simulations that the power of the conventional unit root tests increases with the time span but not with sampling frequency. More important, bringing unit root tests to the high-frequency data is nontrivial. There are many stylized facts of high-frequency finance data, namely, jumps (Andersen et al., 2007a; Lee and Mykland, 2008), conditional heteroskedasticity (Engle, 1982; Bollerslev, 1986; Taylor, 1994), microstructure noise (Aït-Sahalia et al., 2005; Ait-Sahalia and Yu, 2009), and intraday periodicity (Taylor and Xu, 1997; Andersen and Bollerslev, 1997), which may potentially affect the performance of the test.

The main focus of this paper is on the effect of jumps on unit root tests. The presence of jumps in high-frequency data has now been widely recognized in the literature.[5] In the empirical application, we identified 149 jumps in the 10-minute Nasdaq log prices around the peak of the dot-com bubble (from May 1999 to June 2000) and 91 jumps from May 2015 to Jan 2016, with their locations displayed in Figure 8. Some of the jumps identified are of a very large magnitude. The occurrence of jumps might be due to macroeconomic news and company-specific announcements such as share buybacks (Bajgrowicz et al., 2015; Lee, 2011). We show both asymptotically and by simulations that ignoring the presence of jumps leads to a severe size distortion for the Dickey-Fuller test (depending on the number, location, and magnitude of jumps).

The proposed procedures take the presence of jumps into consideration by including a set of dummies in the regression model. We provide the limiting distributions of two test statistics, denoted by DF^J and $DF(J)$, under both the null of a random walk and the alternative of mean reversion or explosiveness (with or without jumps). We consider two versions of these tests: an unfeasible one that relies on the true jump occurrences and a feasible one that relies on a test to identify jumps. We show the asymptotic equivalence of the unfeasible and feasible tests.

In the simulations, we show that in the presence of jumps, the conventional unit root test (which ignores jumps) is undersized for the mean reversion alternative and oversized for the alternative of an explosive process. In contrast, the new tests, which account for the presence of jumps, have satisfactory finite sample performance. The empirical sizes are close to the nominal sizes, while the powers of the new tests are reasonably high even for very small deviations from the random walk. Moreover, the presence of conditional heteroskedasticity, intraday periodicity in volatility and microstructure noise does not affect the performance of the tests when the test window is one quarter or longer and the sampling frequency is 10 minutes or lower. Furthermore, we show that the

[5] See Mancini (2011) for a review on jumps in high-frequency financial data.

right-sided $DF(J)$ test has a high probability of rejecting the null against the explosive alternative when the process is stationary (but very close to a random walk) and N is relatively low. Therefore, we recommend the use of the DF^J test for empirical applications.

Finally, we apply the $DF^{\hat{J}}$ test (a feasible version of DF^J), along with the conventional unit root test, to 10-minute log prices of the Nasdaq composite index around the peak of the dot-com bubble (1999-2000) and the 2015-2016 sell-off periods. We find cases where different conclusions are drawn from DF and $DF^{\hat{J}}$. We attribute these differences to the lack of power of the left-sided DF test and the oversize of the right-sided DF test when jumps are ignored. Moreover, there are several interesting empirical findings. First, we find evidence of deviations from the random walk hypothesis to the explosive direction in late 1999 and to the stationary direction in late 2015. Second, our findings show that the dynamic of the log Nasdaq price switches back to a random walk (from being explosive) as it approaches the peak of the bubble episode. This finding suggests that the $DF^{\hat{J}}$ test could potentially enable investors to withdraw from the market before it collapses. Third, while the dot-com bubble bursts in a random walk fashion, the stock market crash in late 2015 follows a mean-reverting pattern. The last finding provides empirical support for the mildly stationary process of Phillips and Shi (2018) and the random drift martingale process of Phillips and Shi (2019) for crashes.

This paper is closely related to the work of Tao et al. (2019), Kim and Park (2019), and Jiang et al. (2018). Tao et al. (2019) propose new tests for the identification of extreme behaviors in asset prices using high-frequency data. Kim and Park (2019) proposes using the conventional unit root tests for the identification of mean-reverting behaviors and applying unit root tests to Lamperti-transformed data series to distinguish stationary and nonstationary processes. Jiang et al. (2018) analyzes the behavior of the KPSS stationarity test in a continuous-time framework. However, none of these papers considers the impact of the high-frequency features of financial data (especially jumps) on test performance.

The remainder of the paper is organized as follows. Section 2 revisits the conventional unit root test with in-fill asymptotics under both the null and the alternatives. Section 3 introduces the new unit root tests for intraday high-frequency data, provides the limiting distributions of the new test statistics under both the null and the alternative, and discusses the jump detection procedure. Monte Carlo simulations are conducted in Section 4. An empirical illustration using the Nasdaq stock index is proposed in Section 5. Section 6 concludes the paper. All proofs are collected in the appendix.

2 Econometric Method

Consider a set of equally spaced data sampled at an interval Δ. The logarithmic price is denoted by $y_{i\Delta}$ with $i = \{1, \ldots, T\}$. The T observations span across $N = T\Delta$ days. The aim is to detect any deviations of $y_{i\Delta}$ from the random walk using intraday data from a fixed time period (N days). Jumps are not considered here but will be introduced in the next section.

2.1 Hypotheses and Model Specifications

The null hypothesis of a unit root is specified as

$$y_{i\Delta} = y_{(i-1)\Delta} + \sigma\sqrt{\Delta}\varepsilon_{i\Delta}, \tag{1}$$

with initial value y_0, where σ is a constant and $\varepsilon_{i\Delta} \overset{i.i.d}{\sim} N(0, 1)$. The alternative hypothesis is

$$y_{i\Delta} = \alpha_0 + \beta_0 y_{(i-1)\Delta} + \lambda_0 \varepsilon_{i\Delta}, \tag{2}$$

where $\alpha_0 = \mu\left(1 - e^{\theta\Delta}\right)$ with μ and θ being constant, $\beta_0 = e^{\theta\Delta}$ and $\lambda_0^2 = \frac{\sigma^2}{2\theta}\left(e^{2\theta\Delta} - 1\right)$. Model (2) is the exact discrete-time solution of the drift-diffusion process

$$dy_t = \theta\left(y_t - \mu\right)dt + \sigma dw_t, \tag{3}$$

where w_t is the standard Wiener process. It reduces to Model (1) when $\theta = 0$. When $\theta \neq 0$, the autoregressive coefficient

$$\beta_0 = 1 + \theta\Delta + O\left(\Delta^2\right)$$

converges to unity at a rate of $\Delta = N/T$. Given that N is fixed, the process is equivalent to the local-to-unity process of Phillips (1987b) in both the explosive (when $\theta > 0$) and mean reversion (when $\theta < 0$) directions.[6]

The regression model used to test the null hypothesis of a unit root includes an intercept and is specified as follows:

$$y_{i\Delta} = \alpha + \beta y_{(i-1)\Delta} + v_{i\Delta}, \tag{4}$$

where $v_{i\Delta}$ is the error term. The Dickey-Fuller statistic is

$$DF = \left(\hat{\beta} - 1\right)\left[\frac{T\sum_{j=1}^{T} y_{j\Delta}^2 - \left(\sum_{j=1}^{T} y_{(j-1)\Delta}\right)^2}{\sum_{j=1}^{T}\left(y_{j\Delta} - \hat{\alpha} - \hat{\beta} y_{(j-1)\Delta}\right)^2}\right]^{1/2}, \tag{5}$$

where $\hat{\alpha}$ and $\hat{\beta}$ represent the OLS estimates of α and β.

Next, we provide the asymptotic properties of the unit root test under both the null and the alternative. The proofs of Lemma 2.1, Lemma 2.2, Theorem 2.1, and Theorem 2.2 are collected in the online supplement.

2.2 Asymptotics Under the Null

Lemma 2.1 *Under the null hypothesis (1), as $\Delta \to 0$ ($T \to \infty$ with N fixed):*

$$(a)\;\; y_{T\Delta} \Longrightarrow \sigma N^{1/2}\left(w_1 + \gamma\right) \equiv \sigma N^{1/2}\Psi_1;$$

$$(b)\;\; T^{-1}\sum_{j=1}^{T} y_{j\Delta} \Longrightarrow \sigma N^{1/2}\left(\int_0^1 w_s ds + \gamma\right) \equiv \sigma N^{1/2}\Psi_2;$$

$$(c)\;\; T^{-1}\sum_{j=1}^{T} y_{j\Delta}^2 \Longrightarrow \sigma^2 N\left(\int_0^1 w_s^2 ds + \gamma^2 + 2\gamma\int_0^1 w_s ds\right) \equiv \sigma^2 N\Psi_3;$$

$$(d)\;\; T^{-1/2}\sum_{j=1}^{T} y_{(j-1)\Delta}\varepsilon_{j\Delta} \Longrightarrow \frac{1}{2}\sigma N^{1/2}\left(w_1^2 + 2\gamma w_1 - 1\right) \equiv \sigma N^{1/2}\Psi_4;$$

with w_s being the standard Wiener process with variance s and $\gamma = \frac{y_0}{N^{1/2}\sigma}$.

[6] Kim and Park (2019) show that for the drift diffusion process (3), (non-)mean reversion is equivalent to (non-)stationarity. Furthermore, they consider a general null recurrent diffusion process and show that even under this general model setting, one could employ unit root tests to identify mean-reverting behaviors. However, mean reversion is not equivalent to stationarity in the general setting. A process can be nonstationary and mean reverting. Testing for stationarity versus nonstationarity can be achieved by employing the Lamperti transformation before conducting unit root tests.

Theorem 2.1 *Under the null hypothesis (1) and with regression model (4), when the sampling interval $\Delta \to 0$ and the time span N is fixed, the DF test statistic*

$$DF \xrightarrow{L} \frac{-\Psi_2 w_1 + \Psi_4}{\left(\Psi_3 - \Psi_2^2\right)^{1/2}} = \frac{\frac{1}{2}\left(w_1^2 - 1\right) - w_1 \int_0^1 w_s ds}{\left[\int_0^1 w_s^2 ds - \left(\int_0^1 w_s ds\right)^2\right]^{1/2}} \equiv \Upsilon_1.$$

The results in Lemma 2.1 are identical to those in Theorem 6.2 of Phillips (1987a). It is repeated here for ease of comparison. Although the asymptotics of the four quantities in Lemma 2.1 depend on the nuisance parameter γ, the test statistic is asymptotically pivotal. Furthermore, the limiting distribution of DF is identical to its long-span asymptotic (see Hamilton, 1994 for a book reference). This is in sharp contrast to the in-fill limits provided by Phillips (1987a) and Perron (1991), which depend on the nuisance parameter γ. This difference arises from the fact that our regression model includes an intercept, whereas there is no intercept in that of Phillips (1987a) and Perron (1991). Suppose that we did not include an intercept in the regression model as in Phillips (1987a) and Perron (1991). Under the null hypothesis of (1), the test statistics

$$DF \xrightarrow{L} \frac{\frac{1}{2}\left(w_1^2 - 1\right) + \gamma w_1}{\left[\gamma^2 + \int_0^1 w_s^2 ds + 2\gamma \int_0^1 w_s ds\right]^{1/2}}.$$

The proof follows directly from that of Lemma 2.1 and is omitted for brevity.

Remark 2.1 *The null specification* (1) *can be generalized to allow for an asymptotically negligible drift such that*

$$y_{i\Delta} = \mu\Delta^{\eta} + y_{(i-1)\Delta} + \sigma\sqrt{\Delta}\varepsilon_{i\Delta},$$

with μ being a constant and $\eta > 1$. The inclusion of the small drift $\mu\Delta^{\eta}$ does not have any impact on the limiting properties of the DF statistic.

2.3 Asymptotics under Local Alternatives

By recursive substitution, Model (2) becomes

$$y_{i\Delta} = \alpha_0 \frac{1 - e^{i\theta\Delta}}{1 - e^{\theta\Delta}} + \lambda_0 \sum_{j=1}^{i} e^{(i-j)\theta\Delta}\varepsilon_{j\Delta} + e^{i\theta\Delta} y_0. \tag{6}$$

The stochastic component converges to an Ornstein-Uhlenbeck (OU) process in the limit, i.e.,

$$T^{-1/2}\sum_{j=1}^{i} e^{(i-j)\theta\Delta}\varepsilon_{\Delta j} \Longrightarrow J_c(r) = \int_0^r \exp\left(c\left(r-s\right)\right) dw_s \text{ with } r = i/T \text{ and } c = \theta N.$$

Lemma 2.2 *Under the alternative of (2), as $\Delta \to 0$ ($T \to \infty$ with N fixed),*

(a) $y_{T\Delta} \Longrightarrow \sigma N^{1/2}\left[\delta\left(1 - e^c\right) + J_c(1) + e^c\gamma\right] \equiv \sigma N^{1/2}\Xi_1$;

(b) $T^{-1}\sum_{i=1}^{T} y_{i\Delta} \Longrightarrow \sigma N^{1/2}\left[\delta + \int_0^1 J_c(r) dr + (\gamma - \delta)\frac{e^c - 1}{c}\right] \equiv \sigma N^{1/2}\Xi_2$;

(c) $T^{-1}\sum_{i=1}^{T} y_{i\Delta}^2 \Longrightarrow \sigma^2 N\left[\delta^2 + 2\delta(\gamma - \delta)\frac{e^c - 1}{c} + (\gamma - \delta)^2\frac{e^{2c} - 1}{2c}\right.$

$$+\int_0^1 J_c(r)^2\,dr + 2\delta\int_0^1 J_c(r)\,dr + 2(\gamma-\delta)\int_0^1 e^{cr}J_c(r)\,dr\Big] \equiv \sigma^2 N\Xi_3;$$

$$(d)\ \ T^{-1/2}\sum_{i=1}^{T} y_{(i-1)\Delta}\varepsilon_{i\Delta} \Longrightarrow \frac{\sigma N^{1/2}}{2}\left(\Xi_1^2-\gamma^2-2c\Xi_3-1+2c\delta\Xi_2\right) \equiv \sigma N^{1/2}\Xi_4$$

with $J_c(r) = \int_0^r \exp(c(r-s))\,dw_s$, $c=\theta N$, $\gamma = \frac{y_0}{N^{1/2}\sigma}$, *and* $\delta = \frac{\mu}{N^{1/2}\sigma}$.

Lemma A.1 of Perron (1991) is a special case of Lemma 2.2 with $\mu = 0$. The results in Lemma 2.2 are identical to those reported in Lemma 8.1 of Zhou and Yu (2015).

Theorem 2.2 *Under the alternative hypothesis (2), as* $\Delta \to 0$ *(*$T \to \infty$ *with* N *fixed), the DF statistic has the following limit distribution:*

$$DF \xrightarrow{L} \frac{-\Xi_2 w_1 + \Xi_4}{\left(\Xi_3-\Xi_2^2\right)^{1/2}} + c\left(\Xi_3-\Xi_2^2\right)^{1/2} \equiv \Upsilon_1^A,$$

where Ξ_2, Ξ_3 *and* Ξ_4 *are defined in Lemma 2.2.*

The limiting distribution Υ_1^A is continuous with respect to θ. It converges to Υ_1 as $\theta \to 0$. The limiting distribution of the DF statistic (and hence the asymptotic power of the test) depends on the model parameters y_0, N, μ, θ, and σ.

Figure 1: The asymptotic distributions (kernel densities) of the DF test statistic under the null Υ_1 and the alternatives Υ_1^A for $N = 20$ and 60. The value of θ ranges from -0.02 to 0.02 with an increment of 0.0001. We set $y_0 = 6.959$, $\mu = 0.0002$ and $\sigma = 0.009$.

(a) $N = 20$ (b) $N = 60$

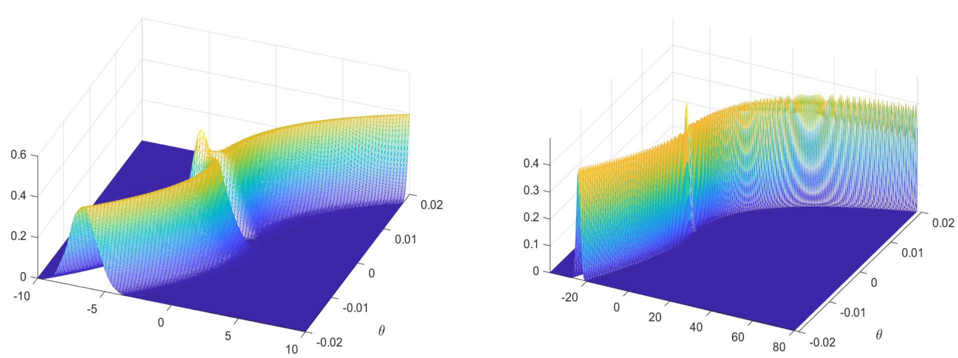


Figure 1 graphs the asymptotic distributions (kernel densities) of the DF statistic for various settings of θ with $N = \{20, 60\}$. We allow θ to take values from -0.02 to 0.02 with an increment of 0.0001. The initial value $y_0 = 6.959$ is the log price of the Nasdaq stock market on January 2, 1996. The parameter $\sigma = 0.009$, which is the average of the estimated spot volatility for the 10-minute Nasdaq log price from January 2, 1996 to December 8, 2017. The time period N equals one month ($N = 20$) and one quarter ($N = 60$). The distributions are obtained from 10,000 replications, approximating the Wiener process by partial sums of standard normal variates with 10,000 steps. The parameter μ is set to 0.0002. The case of $\theta = 0$ corresponds to the null distribution Υ_1.

First, Υ_1^A is a nonlinear function of θ. When $N = 20$, the shape resembles a 'swimmer' with the peak of the limiting distribution of $\theta = 0$ being the head, those of $\theta > 0$ (resp. $\theta < 0$) forming the right (resp. left) shoulder and arm. The right shoulder and arm are always behind the head, reaching back. That is, when $\theta > 0$, the distribution moves sequentially to the right as θ increases, implying a rising power of the right-tailed unit root test. When $\theta < 0$, the limiting distribution changes in a nonmonotonic fashion. It moves first to the right of the null distribution and then gradually to the left as θ deviates further away from zero. The nonmonotonicity occurs when θ is very close to zero (viz. the left shoulder). This implies that when $\theta < 0$ but very close to zero, the right-tailed unit root test rejects the null hypothesis too often (finding evidence of spurious explosive dynamics), while the left-tailed unit root test has no power against $\theta < 0$. This problem disappears when the left arm is ahead of the head (i.e., θ moves further away from zero). Furthermore, from panel (b), the distribution of the DF statistic moves rapidly to the right (resp. left) for positive (resp. negative) θs as the time period N rises to 60. The rate of divergence is faster on the right than on the left.

Remark 2.2 *Under the data generating process (2) and a double asymptotic scheme ($N \to \infty$ and $\Delta \to 0$),*

$$DF \sim \begin{cases} e^{\theta N}\sqrt{\frac{\theta}{2}}\frac{1}{\sigma}\left|y_0 - \mu + \sigma Z_1\right| \to +\infty & \text{if } \theta > 0 \\ N^{1/2}\frac{\theta}{\sigma}\left(-\frac{1}{2\theta}\right)^{1/2} \to -\infty & \text{if } \theta < 0 \end{cases},$$

where $Z_1 \sim N\left(0, \frac{1}{2\theta}\right)$. The proof follows directly from the results of Wang and Yu (2016). See Appendix B for details. The DF statistic diverges to positive and negative infinity when $\theta > 0$ and $\theta < 0$, respectively. This result suggests that one could obtain the asymptotic consistency of the test by allowing the time period N go to infinity. Furthermore, when $\theta > 0$, the divergence rate is exponential, i.e., $O_p\left(e^{\theta N}\right)$. The divergence rate of the DF statistic is $O_p\left(N^{1/2}\right)$ when $\theta < 0$, which is slower than that of DF when $\theta > 0$. This result is consistent with our observation from Figure 1.

3 Unit Root Tests for High-frequency Data

As highlighted in Bauwens et al. (2012), empirical studies have shown that stochastic diffusion models driven by Brownian motion fail to explain some characteristics of asset returns. One of the most important features of financial assets is the presence of discontinuities in prices, also called jumps. See Andersen et al. (2007a) and Lee and Mykland, 2008, among others. Several jump-diffusion processes have been proposed in the literature to account for the presence of either small (infinite activity) jumps or large finite activity jumps. See, for example, Merton (1976); Ahn and Thompson (1988); Kou (2002); Mancini (2011).

Here, jumps in log prices are additive and governed by a compound Poisson process J_t. The log prices

$$dy_t = \theta\left(y_t - \mu\right)dt + \sigma dw_t + dJ_t. \tag{7}$$

The exact discrete time solution of (7) is

$$y_{i\Delta} = \alpha_0 + \beta_0 y_{(i-1)\Delta} + \lambda_0\varepsilon_{i\Delta} + \sum_{k=K_{(i-1)\Delta}}^{K_{i\Delta}} e^{\theta(i\Delta - \tau_k)}\xi_k, \tag{8}$$

where α_0, β_0, and λ_0 are identical to those in (2), $K_{i\Delta}$ is the total number of jumps within the interval $[0, i\Delta]$ and follows a Poisson process with intensity λ, and τ_k is the location of the k^{th} jump. The jump size $\{\xi_k\}$ is a sequence of independent random variables governed by law f, e.g., the lognormal distribution (Merton, 1976) or double exponential distribution (Kou, 2002). Although the jump component is stochastic under (7), it is sufficient for the purposes of this paper to assume that the number, locations, and sizes of the jumps are deterministic. This assumption is analogous to that for bubble generating processes. While Evans (1991) consider a data generating process where the collapse of bubbles is governed by a Bernoulli process, Phillips et al. (2011) and Phillips et al. (2015a,b) assume deterministic switching points of bubbles for their analysis of bubble origination and termination dates.

Assume there are K jumps within the sample period. The locations and magnitudes of the jumps are denoted by $\tau_k = \lfloor r_k T \rfloor$ and ϕ_k, respectively, with $k = 1, \ldots, K$ and $\lfloor . \rfloor$ signifying the integer of the argument. The jump dummy $I_{i\Delta}^k$ takes value one at period τ_k when the k^{th} jump occurs and zero otherwise so that $\sum_{i=1}^T I_{i\Delta}^k = 1$ (for $k = 1, \ldots, K$). The null hypothesis of a unit root with jumps is

$$y_{i\Delta} = \sum_{k=1}^{K} \phi_k I_{i\Delta}^k + y_{(i-1)\Delta} + \sigma\sqrt{\Delta}\varepsilon_{i\Delta}, \tag{9}$$

while under the alternative,

$$y_{i\Delta} = \alpha_0 + \sum_{k=1}^{K} \phi_k I_{i\Delta}^k + \beta_0 y_{(i-1)\Delta} + \lambda_0 \varepsilon_{i\Delta}. \tag{10}$$

In the special case where the sizes of the jumps are identical, i.e., $\phi_1 = \cdots = \phi_K = \phi$, the jump component

$$\sum_{k=1}^{K} \phi_k I_{i\Delta}^k = \phi \sum_{k=1}^{K} I_{i\Delta}^k = \phi I_{i\Delta}^* \text{ with } I_{i\Delta}^* = \sum_{k=1}^{K} I_{i\Delta}^k.$$

The dummy variable $I_{i\Delta}^*$ takes value one when there is a jump and zero otherwise (i.e., $\sum_{i=1}^T I_{i\Delta}^* = K$). By splitting the jump indicator $I_{i\Delta}^*$ into k orthogonal variables $\left\{I_{i\Delta}^k\right\}_{k=1}^K$, we therefore allow for different jump sizes.

We provide the limiting properties of $y_{T\Delta}$, $\sum_{i=1}^T y_{i\Delta}$, $\sum_{i=1}^T y_{i\Delta}^2$, and $\sum_{i=1}^T y_{(i-1)\Delta}\varepsilon_{i\Delta}$ under the null and the alternative in Lemma 3.1 and 3.2, respectively. The proofs of these two lemmas are collected in the appendix.

Lemma 3.1 *Under the null hypothesis (9), as $\Delta \to 0$ ($T \to \infty$ with N fixed):*

$$(a)\ \ y_{T\Delta} \Longrightarrow \sigma N^{1/2} \left(\Psi_1 + \sum_{k=1}^{K} \zeta_k \right) \equiv \sigma N^{1/2} \tilde{\Psi}_1;$$

$$(b)\ \ T^{-1} \sum_{i=1}^{T} y_{i\Delta} \Longrightarrow \sigma N^{1/2} \left[\Psi_2 + \sum_{k=1}^{K} \zeta_k \left(1 - r_k\right) \right] \equiv \sigma N^{1/2} \tilde{\Psi}_2;$$

$$(c)\ \ T^{-1} \sum_{i=1}^{T} y_{i\Delta}^2 \Longrightarrow \sigma^2 N \left[\Psi_3 + \Delta_1 + 2 \sum_{k=1}^{K} \zeta_k \int_{r_k}^{1} w_s ds + 2\gamma \sum_{k=1}^{K} \zeta_k \left(1 - r_k\right) \right] \equiv \sigma^2 N \tilde{\Psi}_3;$$

$$(d)\ \ T^{-1/2} \sum_{i=1}^{T} y_{(i-1)\Delta}\varepsilon_{i\Delta} \Longrightarrow \frac{\sigma N^{1/2}}{2} \left[\tilde{\Psi}_1^2 - \gamma^2 - 1 - \sum_{k=1}^{K} \zeta_k^2 - 2 \sum_{k=1}^{K} \zeta_k \left(w_{r_k} + \gamma + \Delta_2 \right) \right] \equiv \sigma N^{1/2} \tilde{\Psi}_4,$$

where w_s is the standard Wiener process with variance $s \in [0,1]$, $\gamma = \frac{y_0}{N^{1/2}\sigma}$, $\zeta_k = \frac{\phi_k}{\sigma N^{1/2}}$,

$$\Delta_1 = \begin{cases} \zeta_1^2 \left(1 - r_1\right) & \text{if } K = 1 \\ \sum_{k=1}^{K-1} \left(r_{k+1} - r_k\right) \left(\sum_{j=1}^{k} \zeta_j\right)^2 + \left(1 - r_K\right) \left(\sum_{j=1}^{K} \zeta_j\right)^2 & \text{if } K > 1 \end{cases},$$

and

$$\Delta_2 = \begin{cases} 0 & \text{if } K = 1 \\ \sum_{j=1}^{k-1} \zeta_j & \text{if } K > 1 \end{cases}.$$

We denote the limiting properties of the above four quantities under the null without jumps (1) by Ψ (with their exact forms in Lemma 2.1) and those under the null with jumps (9) by $\tilde{\Psi}$. It is evident from Lemma 3.1 that the jump component $\sum_{k=1}^{K} \phi_k I_{i\Delta}^k$ has an asymptotic impact on those four quantities. There are additional terms in $\tilde{\Psi}$, relating to the number of jumps K, the jump sizes ϕ_k, and the (fractional) location of the jumps r_k.[7]

Lemma 3.2 *Under the alternative model (10), as $\Delta \to 0$ ($T \to \infty$ with N fixed):*

$$(a)\ y_{T\Delta} \Longrightarrow \sigma N^{1/2} \left[\Xi_1 + \sum_{k=1}^{K} \zeta_k e^{(1-r_k)c} \right] \equiv \sigma N^{1/2} \tilde{\Xi}_1;$$

$$(b)\ T^{-1} \sum_{i=1}^{T} y_{i\Delta} \Longrightarrow \sigma N^{1/2} \left\{ \Xi_2 + \frac{1}{c} \sum_{k=1}^{K} \zeta_k \left[e^{(1-r_k)c} - 1 \right] \right\} \equiv \sigma N^{1/2} \tilde{\Xi}_2;$$

$$(c)\ T^{-1} \sum_{i=1}^{T} y_{i\Delta}^2 \Longrightarrow \sigma^2 N \left\{ \Xi_3 + \Delta_3 + \frac{1}{c} \delta \sum_{k=1}^{K} \zeta_k \left[2e^{c(1-r_k)} - 2 - e^{c(2-r_k)} + e^{r_k c} \right] \right.$$
$$\left. + 2 \sum_{k=1}^{K} \zeta_k \int_{r_k}^{1} e^{c(r-r_k)} J_c(r)\, dr + \frac{1}{c} \gamma \sum_{k=1}^{K} \zeta_k e^{r_k c} \left[e^{2c(1-r_k)} - 1 \right] \right\} \equiv \sigma^2 N \tilde{\Xi}_3;$$

$$(d)\ T^{-1/2} \sum_{i=1}^{T} y_{(i-1)\Delta} \varepsilon_{i\Delta} \Longrightarrow \frac{\sigma N^{1/2}}{2} \left\{ \tilde{\Xi}_1^2 - \gamma^2 - 2c\tilde{\Xi}_3 - 1 + 2c\delta\tilde{\Xi}_2 - \sum_{k=1}^{K} \zeta_k^2 \right.$$
$$\left. -2 \sum_{k=1}^{K} \zeta_k \left[\delta \left(1 - e^{r_k c}\right) + J_c(r_k) + e^{r_k c} \gamma + \Delta_4 \right] \right\} \equiv \sigma N^{1/2} \tilde{\Xi}_4.$$

where $c = \theta N$, $\delta = \frac{\mu}{N^{1/2}\sigma}$, and

$$\Delta_3 = \begin{cases} \zeta_1^2 \frac{e^{2c(1-r_1)}-1}{2c} & \text{if } K = 1 \\ \sum_{k=1}^{K-1} \left(\sum_{j=1}^{k} \zeta_j e^{-r_j \theta} \right)^2 e^{2r_k \theta} \frac{e^{2\theta(r_{k+1}-r_k)}-1}{2c} + \left(\sum_{j=1}^{K} \zeta_j e^{-r_j \theta} \right)^2 e^{2r_K \theta} \frac{e^{2\theta(1-r_K)-1}}{2c} & \text{if } K > 1 \end{cases},$$

$$\Delta_4 = \begin{cases} 0 & \text{if } K = 1 \\ \sum_{j=1}^{k-1} e^{(r_k - r_j)\theta} \zeta_j & \text{if } K > 1 \end{cases}.$$

Analogously, we use Ξ to denote limiting properties under (2) and $\tilde{\Xi}$ for those under (10). As in Lemma 3.1, the impact of the jump component $\sum_{k=1}^{K} \phi_k I_{i\Delta}^k$ under the alternative does not disappear in the limit. We observe additional terms in $\tilde{\Xi}$, which depend on the three jump characteristics (i.e., K, ϕ_K, and r_k) as well as θ and N through the parameter c.

3.1 Unit Root Test with Known Jump Location

Assume for now that the number of jumps and their locations are known. This assumption will be relaxed in Section 3.2. The regression model used to test the null hypothesis of a random walk with

[7] We assume that there is a finite number of jumps and that the magnitude of the jumps is finite. One could potentially relax this assumption to allow for an infinite number of jumps K or the magnitude of jumps ϕ_k to go to infinity. We leave this extension to future research.

jumps is

$$y_{i\Delta} = \alpha + \sum_{k=1}^{K} \phi_k I_{i\Delta}^k + \beta y_{(i-1)\Delta} + v_{i\Delta}. \tag{11}$$

We use the notation $\tilde{}$ to denote the OLS estimates of the model coefficients in (11). The corresponding test statistic, denoted DF^J, is

$$DF^J = \left(\tilde{\beta} - 1\right) \left[\frac{T \sum_{i=1}^{T} y_{i\Delta}^2 - \left(\sum_{i=1}^{T} y_{(i-1)\Delta}\right)^2}{\sum_{i=1}^{T} \left(y_{i\Delta} - \tilde{\beta} y_{(i-1)\Delta} - \tilde{\alpha} - \sum_{k=1}^{K} \tilde{\phi}_k I_{i\Delta}^k\right)^2} \right]^{1/2}.$$

The limiting distributions of the DF^J statistic under both the null and the alternatives are provided below. Moreover, we show the limiting distributions of the DF statistic under the null of (9) and the alternative of (10) (i.e., one ignores the presence of jumps and applies the standard unit root test).

3.1.1 Asymptotics of the DF^J Statistic

Theorem 3.1 *Under the null hypothesis of a random walk with jumps* (9)*, the OLS estimators have the following limiting properties: as* $\Delta \to 0$ *(*$T \to \infty$ *with* N *fixed),*

$$\begin{aligned} T^{1/2}\left(\tilde{\phi}_k - \phi_k\right) &\implies N(0, \sigma^2 N); \\ \tilde{\sigma}_v^2 &\to \sigma^2 N, \end{aligned}$$

where $\tilde{\sigma}_v^2 = \sum_{i=1}^{T} \left(y_{i\Delta} - \tilde{\beta} y_{(i-1)\Delta} - \tilde{\alpha} - \sum_{k=1}^{K} \tilde{\phi}_k I_{i\Delta}^k\right)^2$*. The test statistic* DF^J *has the following limiting distribution*

$$DF^J \implies \frac{-\tilde{\Psi}_2 w_1 + \tilde{\Psi}_4}{\left(\tilde{\Psi}_3 - \tilde{\Psi}_2^2\right)^{1/2}} \equiv \Upsilon_2. \tag{12}$$

The proof of this theorem is in Appendix C. Theorem 3.1 suggests that given the exact number of jumps and their location, the estimated jump sizes $\tilde{\phi}_k$ and the error variance $\tilde{\sigma}_v^2$ are consistent. Furthermore, the limiting distribution of DF^J under the new setting (with jumps) is very different from Υ_1 (i.e., in the absence of jumps, see Theorem 2.1). The numerator of Υ_2 can be rewritten as

$$\left[\frac{1}{2}\left(w_1^2 - 1\right) - w_1 \int_0^1 w_s ds\right] + \frac{1}{2}\left[\left(\sum_{k=1}^{K} \zeta_k\right)^2 - \sum_{k=1}^{K} \zeta_k^2\right] - \sum_{k=1}^{K} \zeta_k \Delta_2,$$

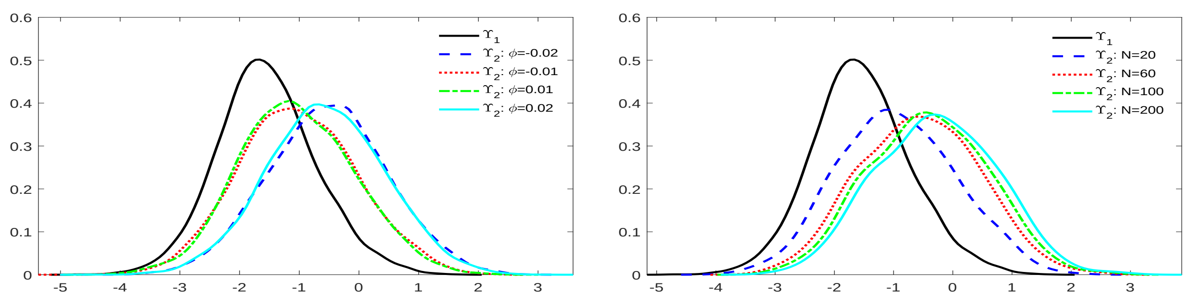
and the denominator of Υ_2 is the square root of the following quantity:

$$\begin{aligned} &\left[\int_0^1 w_s^2 ds - \left(\int_0^1 w_s ds\right)^2\right] + \left[\Delta_1 + (1 - r_K)\left(\sum_{j=1}^{K} \zeta_j\right)^2 - \left(\sum_{k=1}^{K} \zeta_k (1 - r_k)\right)^2\right] \\ &+2\left[\sum_{k=1}^{K} \zeta_k \int_{r_k}^1 w_s ds - \int_0^1 w_s ds \sum_{k=1}^{K} \zeta_k (1 - r_k)\right]. \end{aligned}$$

The limiting distribution of DF^J does not depend on y_0 but on parameters related to jumps, i.e., r_k and ζ_k (including ϕ_k, σ, and N). Next, we simulate the distribution Υ_2. For simplicity, we assume one jump per week such that $K = \lfloor N/5 \rfloor + 1$, $\tau_1 = 1/\Delta$, and $\tau_k = 5(k-1)/\Delta$ for $k > 1$. The sign and magnitude of the jumps are assumed to be the same (i.e., $\phi_k = \phi$). We set $\phi = \{-0.02, -0.01, 0, 0.01, 0.02\}$ and $N = 60$ in the left panel and $N = \{20, 60, 100, 200\}$ and $\phi = 0.02$ in the right panel. We set $\sigma = 0.009$ as in Section 2.3. From Figure 2, Υ_2 is always on the right of Υ_1 and shifts to the right as the magnitude of the jumps increases (regardless of the sign of the jumps) or the time period N expands.

Figure 2: The asymptotic distribution of DF^J (kernel densities) under the null hypothesis of a random walk with jumps. We assume $K = \lfloor N/5 \rfloor + 1$, $\tau_1 = 1/\Delta$, $\tau_k = 5(k-1)/\Delta$ for $k > 1$, and $\phi_k = \phi$. The jump size $\phi = \{-0.02, -0.01, 0.01, 0.02\}$ with $N = 60$ in the left panel and $N = \{20, 60, 100, 200\}$ with $\phi = 0.02$ in the right panel.

(a) $N = 60$ (b) $\phi = 0.02$



Theorem 3.2 *Under the alternative hypothesis (10), the limiting properties of the OLS estimators are as follows:*

$$T^{1/2}\left(\tilde{\phi}_k - \phi_k\right) \Rightarrow N(0, \sigma^2 N),$$
$$\tilde{\sigma}_v^2 \to \sigma^2 N.$$

The test statistic DF^J has the limiting distribution of

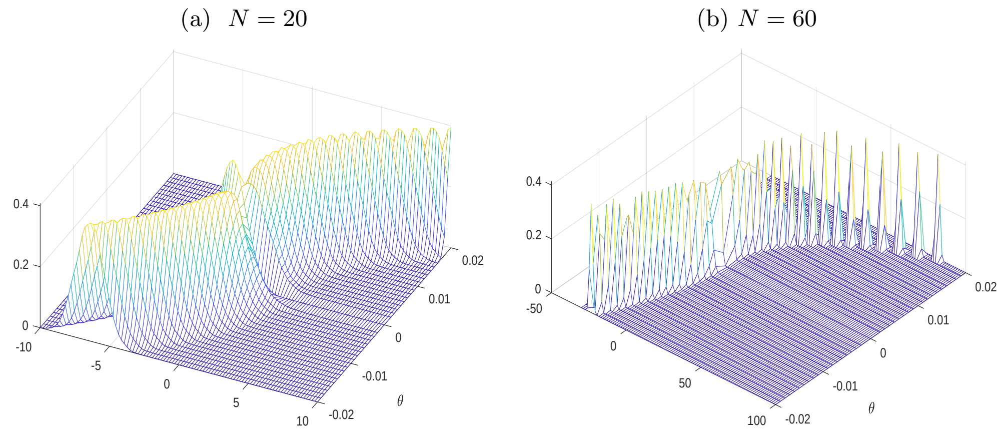
$$DF^J \Longrightarrow \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1}{\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2}} + c\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2} \equiv \Upsilon_2^A. \tag{13}$$

Theorem 3.2 shows that under the alternative (10), the OLS estimators $\tilde{\phi}_k$ and $\tilde{\sigma}_\varepsilon^2$ are consistent. The limiting distribution of DF^J depends on θ and N through the parameter c, in addition to the nuisance parameters in Υ_2. We now plot the asymptotic distribution of DF^J against θ with $N = \{20, 60\}$ in Figure 3. To reduce computation, we allow θ to vary from -0.002 to -0.02 on the left and from 0.002 to 0.02 on the right, with an increment of 0.001 (instead of 0.0001 for Figure 1). The setting of jumps is identical to that in Figure 2. The other parameters are the same as those in Figure 1. One can see that the pattern of the DF^J distribution is similar to that of the DF statistic in Figure 1.

For practical implementation, one needs to estimate r_k, ϕ_k and σ before simulating the asymptotic critical values. As shown in Theorem 3.1 and Theorem 3.2, given the locations of the jumps, the magnitude of the jumps ϕ_k can be consistently estimated by OLS with Equation (11), while σ^2 can be consistently estimated as

$$\tilde{\sigma}^2 = \frac{1}{N}\sum_{i=1}^{T}\left(y_{i\Delta} - \tilde{\beta} y_{(i-1)\Delta} - \tilde{\alpha} - \sum_{k=1}^{K}\tilde{\phi}_k I_{i\Delta}^k\right)^2 \to \sigma^2$$

Figure 3: The asymptotic distributions (kernel densities) of the DF^J test statistic under the null and the alternative when $N = 20$ and 60. The value of θ ranges from -0.02 to -0.002 on the left and from 0.002 to 0.02 on the right, with an increment of 0.001. We set $y_0 = 6.959$, $\mu = 0.0002$ and $\sigma = 0.009$, $K = \lfloor N/5 \rfloor + 1$, $\tau_1 = 1/\Delta$, $\tau_k = 5(k-1)/\Delta$ for $k > 1$, and $\phi_k = \phi = 0.02$.



under both the null and the alternative. The time period N is known for a given sample. The location of jumps r_k can be identified by the procedure introduced in Section 3.2.1.

3.1.2 Asymptotic of the DF Statistic in the Presence of Jumps

We first show that the standard unit root test, which compares the DF test statistic with critical values obtained from Υ_1, has incorrect size in the presence of additive jumps.

Theorem 3.3 *Suppose that the data generating process is* (9) *and that one ignores the presence of jumps by estimating Model* (4)*. As $\Delta \to 0$ ($T \to \infty$ with N fixed):*

$$\hat{\sigma}_v^2 = \sum \left(y_{j\Delta} - \hat{\beta} y_{(j-1)\Delta} - \hat{\alpha} \right)^2 \to \sigma^2 N \left(1 + \sum_{k=1}^{K} \zeta_k^2 \right).$$

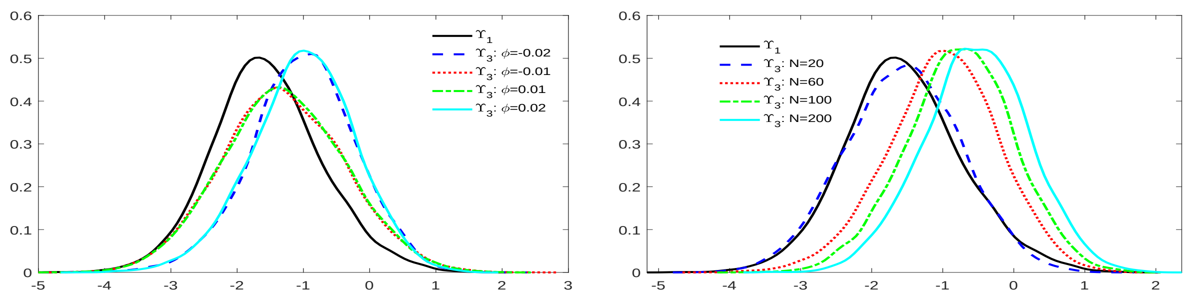
The DF test statistic has the following limiting distribution:

$$DF \Longrightarrow \frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1 + \sum_{k=1}^{K} \zeta_k \left(w_{r_k} + \gamma + \Delta_2 - \tilde{\Psi}_2 \right)}{\left(1 + \sum_{k=1}^{K} \zeta_k^2 \right)^{1/2} \left(\tilde{\Psi}_3 - \tilde{\Psi}_2^2 \right)^{1/2}} \equiv \Upsilon_3. \tag{14}$$

It is evident from Theorem 3.3 that the estimated model error variance $\hat{\sigma}_v^2$ is inconsistent. The limiting distribution of the DF statistic under (9) is Υ_3, instead of Υ_1. The unit root tests that compare the DF test statistic with critical values from Υ_1 will, therefore, have size distortions. In Figure 4, we simulate Υ_3 and compare it with Υ_1. As before, we set $y_0 = 6.959$, $\sigma = 0.009$, $\phi = \{-0.02, -0.01, 0.01, 0.02\}$ and $N = 60$ in the left panel and $N = \{20, 60, 100, 200\}$ and $\phi = 0.02$ in the right panel. One can see that the distribution of the DF statistic Υ_3 moves to the right of Υ_1 when jumps (both positive and negative) are ignored. This implies that the left-sided DF test is undersized in the presence of jumps, while the right-sided DF test is oversized in the presence of jumps.

Figure 4: The asymptotic distributions (kernel densities) of DF under the null hypothesis of a random walk with jumps. We assume $K = \lfloor N/5 \rfloor + 1$, $\tau_1 = 1/\Delta$, $\tau_k = 5(k-1)/\Delta$ for $k > 1$, and $\phi_k = \phi$. The jump size $\phi = \{-0.02, -0.01, 0.01, 0.02\}$ with $N = 60$ in the left panel and $N = \{20, 60, 100, 200\}$ with $\phi = 0.02$ in the right panel.

(a) $N = 60$ (b) $\phi = 0.02$



As correctly pointed out by an anonymous referee, one can obtain the correct size for a test using the DF statistic by constructing critical values from Υ_3 in the presence of jumps, provided consistent estimates of the nuisance parameters and knowledge of jump occurrence. Theorem 3.4 provides the limiting properties of the DF statistic under the alternative with jumps (10).

Theorem 3.4 *Suppose that the data generating process is* (10) *and the regression model is (4). As* $\Delta \to 0$ *($T \to \infty$ with N fixed):*

$$\hat{\sigma}_v^2 = \sum \left(y_{j\Delta} - \hat{\beta} y_{(j-1)\Delta} - \hat{\alpha} \right)^2 \to \sigma^2 N \left(1 + \sum_{k=1}^{K} \zeta_k^2 \right).$$

The DF test statistic has the following limiting distribution:

$$\begin{aligned} DF \quad &\Longrightarrow \quad \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1 + \sum_{k=1}^{K} \varsigma_k \left[\delta \left(1 - e^{r_k c}\right) + J_c\left(r_k\right) + e^{r_k c} \gamma + \Delta_4 - \tilde{\Xi}_2 \right]}{\left(1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}\right)^{1/2} \left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2}} + c \left(\frac{\tilde{\Xi}_3 - \tilde{\Xi}_2^2}{1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}} \right)^{1/2} \\ &\equiv \quad \Upsilon_3^A. \end{aligned}$$

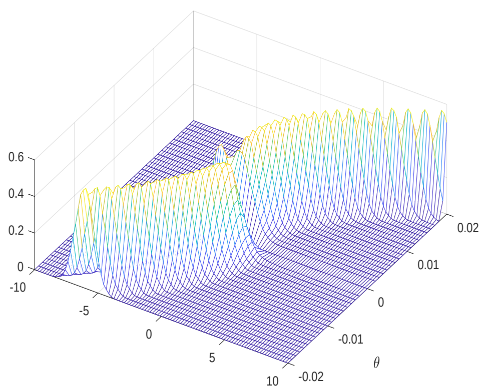
It is obvious that σ cannot be consistently estimated with Regression (4) in the presence of jumps. To compute critical values from Υ_3, one would need to estimate the nuisance parameters from regression (11) as for the DF^J test. We denote the test based on the DF statistic and distribution Υ_3 by $DF(J)$. To avoid confusion, we refer to the DF statistic of the $DF(J)$ test as the $DF(J)$ statistic and that of the standard DF test as the DF statistic.

The asymptotic distribution of the $DF(J)$ statistic is presented in Figure 5. The shapes of the density functions are similar to those of DF and DF^J in Figures 1 and 3. In particular, when θ takes negative but small values, the distribution moves to the right (instead of the left) of the null distribution. As such, the right-sided unit root test might falsely reject the null of unit root against explosiveness when the true θ is negative but very close to zero.

As an illustration, let us consider the case of $\theta = -0.002$. Figure 6 displays the limiting distributions of each test statistic (DF, DF^J and $DF(J)$) under the null and the alternative of $\theta = -0.002$ with $N = \{20, 60\}$. When $N = 20$, the alternative distribution of all three statistics moves to the right of their null distribution. The movement of DF^J is the least significant, implying

Figure 5: The asymptotic distributions (kernel densities) of the $DF(J)$ test statistic under the null and the alternative when $N = 20$ and 60. The value of θ ranges from -0.02 to -0.002 on the left of $\theta = 0$ and from 0.002 to 0.02 on the right, with an increment of 0.001. We set $y_0 = 6.959$, $\mu = 0.0002$ and $\sigma = 0.009$, $K = \lfloor N/5 \rfloor + 1$, $\tau_1 = 1/\Delta$, $\tau_k = 5(k-1)/\Delta$ for $k > 1$, and $\phi_k = \phi = 0.02$.

(a) $N = 20$



(b) $N = 60$

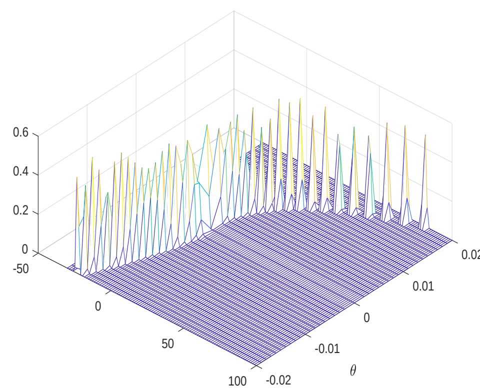
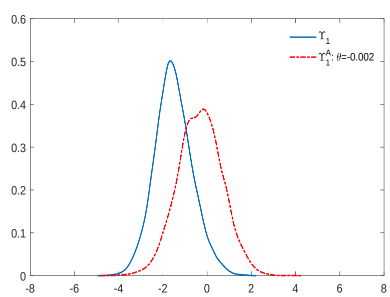
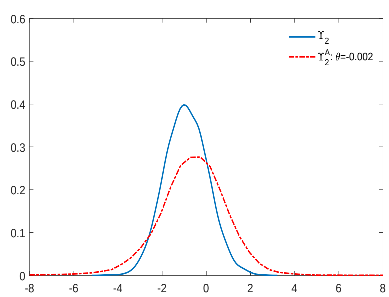


Figure 6: The asymptotic null and alternative distributions of the DF, DF^J, and $DF(J)$ test statistics when $\theta = -0.002$ and $N = \{20, 60\}$.

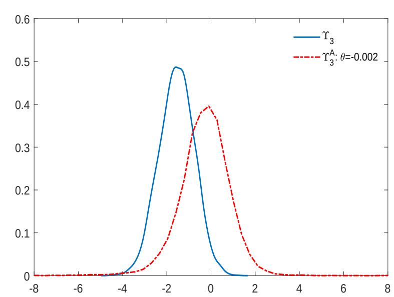
(a) DF and $N = 20$



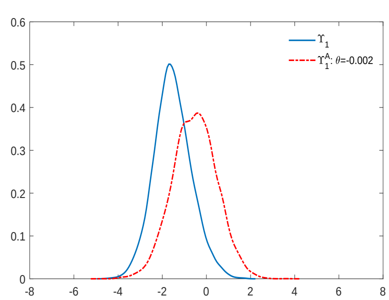
(b) DF^J and $N = 20$



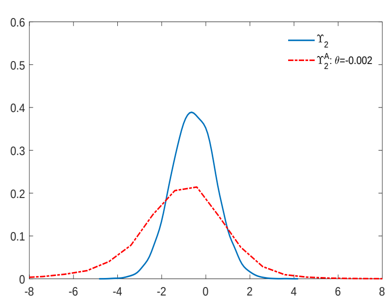
(c) $DF(J)$ and $N = 20$



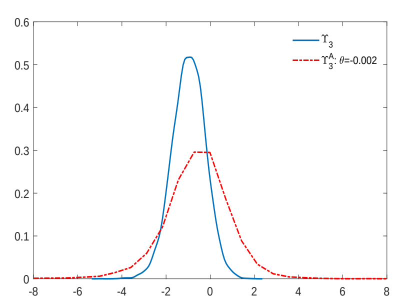
(d) DF and $N = 60$



(e) DF^J and $N = 60$



(f) $DF(J)$ and $N = 60$



the lowest false rejection probability of the right-tailed DF^J test. As the time period increases to 60, the right movements of DF and $DF(J)$ decrease, whereas the DF^J distribution moves to the left instead. This suggests that unlike the other two methods, the false identification issue of the DF^J test disappears when N increases to 60. The DF^J has a clear advantage over $DF(J)$ in this regard. Consequently, we will focus on DF^J rather than $DF(J)$ to account for jumps in the empirical application and advocate not considering small windows (i.e., $N \leq 20$) with one-sided tests to avoid systematically wrong conclusions.

3.2 Unit Root Test with Unknown Jump Location

We consider a two-step procedure for the case with unknown jumps. The first step is to identify the locations of the jumps $\hat{\tau}_k$ and, therefore, also the number of jumps $\hat{K}$ to construct the jump dummies $\hat{I}_{i\Delta}^k$. The method employed to identify the jumps and the limiting properties of the jump test are discussed in Section 3.2.1. In the second step, we conduct the DF^J test by replacing K and $I_{i\Delta}^k$ in regression (11) with $\hat{K}$ and $\hat{I}_{i\Delta}^k$. The limiting properties of the feasible version of the DF^J test, denoted by $DF^{\hat{J}}$, are discussed in Section 3.2.2.

3.2.1 Jump Identification

The most popular approach to the estimation of jump arrival time(s) is probably that proposed by Lee and Mykland (2008, LM hereafter). The data generating process under the null hypothesis of no jumps, considered by Lee and Mykland (2008), is

$$dy_t = \mu_t dt + \sigma_t dw_t, \tag{15}$$

where the drift and diffusion coefficients μ_t and σ_t are assumed not to change dramatically over a short time interval. See Lee and Mykland (2008) for further details on the assumptions. This model includes (3) as a special case with $\mu_t = \theta\,(y_t - \mu)$ and $\sigma_t = \sigma$. Under the alternative,

$$dy_t = \mu_t dt + \sigma_t dw_t + x_t d\tilde{J}_t, \tag{16}$$

where $\tilde{J}_t$ is a counting process independent of w_t and x_t is the jump size, which is assumed to be predictable.

However, the LM jump test suffers from a significant downward size distortion and has low power in finite samples when the drift coefficient μ_t is large in size, as shown by Laurent and Shi (2020). An example given by Laurent and Shi (2020) is the drift diffusion process (3) with θ and hence the drift coefficient $\mu_t = \theta\,(y_t - \mu)$ being nonzero. They propose a simple modification to the Lee and Mykland (2008) test and show a dramatic improvement in test performance. As shown in our empirical application, deviations from the random walk (i.e., nonzero θ) are not rare events. It is, therefore, important to account for such a feature and rely on the jump identification procedure of Laurent and Shi (2020), which is less sensitive to μ_t.

Let $r_{i\Delta} = y_{i\Delta} - y_{(i-1)\Delta}$ denote the log return at time $i\Delta$, $\hat{m}_{i\Delta}$ be the median of the past M log returns (prior to and including the current observation), and $r_{i\Delta}^* = r_{i\Delta} - \hat{m}_{i\Delta}$ be the centered log return. The test statistic of Laurent and Shi (2020), denoted by $U_{i\Delta}$, is constructed from the centered log returns (instead of the raw return $r_{i\Delta}$ as in Lee and Mykland, 2008) such that

$$U_{i\Delta} = \frac{r_{i\Delta}^*}{\hat{\sigma}_{i\Delta}^*} \text{ with } \hat{\sigma}_{i\Delta}^* = \sqrt{\frac{1}{M} BV_{i\Delta}^*}, \tag{17}$$

where $BV_{i\Delta}^* = \frac{\pi}{2}\frac{M}{M-1}\sum_{j=i-M+2}^{i} |r_{j\Delta}^*||r_{(j-1\Delta)}^*|$ is the bipower variation computed on centered log returns. Under the null hypothesis, the test statistic $U_{i\Delta}$ follows a standard normal distribution Z. As in Lee and Mykland (2008), we reject the null hypothesis of no jump at period $i\Delta$ when $|U_{i\Delta}| > cv_{L,\alpha_L}$, where

$$cv_{L,\alpha_L} = C_L + S_L \beta_L, \tag{18}$$

with $C_L = (2\log L)^{1/2} - \frac{1}{2}(2\log L)^{-1/2}[\log 4\pi + \log(\log L)]$, $\beta_L = -\log\left[-\log\left(1 - \alpha_L/2\right)\right]$, and $S_L = (2\log L)^{-1/2}$, L being the number of tests conducted. The critical value is derived from extreme value theory for the purpose of controlling for the oversize issue of multiple tests. In all our simulations and the empirical application, we set L to the total number of observations per month (of 20 days) and α_L to 0.75 so that the expected number of spurious detected jumps is one every four months.

Remark 3.1 *Assume that $L \to \infty$ at the same rate of T and $\beta_L \to \infty$ at a rate that is slower than $\sqrt{L \log(L)}$. It follows that*

$$cv_{L,\alpha_L} = C_L + S_L \beta_L = C_L [1 + o_p(1)] = O\left(\sqrt{2 \log L}\right) = O\left(\sqrt{-2 \log(\Delta)}\right).$$

Furthermore, under the assumption that $M = O(\Delta^a)$ with $-1 < a < -1/2$, we have the asymptotic equivalence of $\hat{\sigma}^*_{i\Delta}$ and $\sigma\sqrt{\Delta}$ (Lee and Mykland, 2008). Suppose that there is a jump at period $i\Delta$. The probability of correctly identifying the jump is

$$\begin{aligned} P(|U_{i\Delta}| > cv_{L,\alpha_L}) &= P\left(\left|\frac{r^*_{i\Delta}}{\hat{\sigma}^*_{i\Delta}}\right| > cv_{L,\alpha_L}\right) = P\left(|r^*_{i\Delta}| > cv_{L,\alpha_L} \hat{\sigma}^*_{i\Delta}\right) \\ &\sim P\left(|r^*_{i\Delta}| > \sigma\sqrt{-2 \log(\Delta)\Delta}\right) = 1 - F_{|r^*_{i\Delta}|}\left(\sigma\sqrt{-2 \log(\Delta)\Delta}\right) \to 1 \end{aligned}$$

as $\Delta \to 0$, which follows the same argument as in Lee (2012). Now, suppose that there is no jump in period $i\Delta$. The probability of not rejecting the null is

$$P(|U_{i\Delta}| \leq cv_{L,\alpha_L}) \sim 1 - 2\Phi\left(-\sqrt{2 \log L}\right) \to 1,$$

as $\Delta \to 0$, where Φ is the cumulative distribution function of the standard normal distribution, given that $U_{i\Delta}$ converges to the standard normal distribution under the null.

Remark 3.2 *The jump dummy*

$$\hat{I}^*_{i\Delta} = \mathbf{1}(|U_{i\Delta}| > cv_{L,\alpha_L}) \to I^*_{i\Delta},$$

and therefore $\hat{I}^k_{i\Delta} \to I^k_{i\Delta}$ as $\Delta \to 0$. This result follows directly from Remark 3.1. Consequently, the estimated number of jumps $\hat{K}$ and their locations $\hat{\tau}_k$ are consistently estimated, i.e.,

$$\hat{K} = \sum_{i=1}^{T} \hat{I}^*_{i\Delta} \to \sum_{i=1}^{T} I^*_{i\Delta} = K \text{ and } \hat{\tau}_k = \arg\max_i \left\{\hat{I}^k_{i\Delta}\right\} \to \tau_k.$$

As in Andersen et al. (2007b), Boudt et al. (2011), and Laurent and Shi (2020), we also take the intraday periodicity in the volatility into consideration. The jump test statistic is

$$U^*_{i\Delta} = \frac{U_{i\Delta}}{\hat{f}^*_{i\Delta}},$$

where $\hat{f}^*_{i\Delta}$ is a robust-to-jumps estimate of the intraday periodicity. In the empirical application, we assume $\hat{f}^*_{i\Delta}$ to be the same across weeks but to vary within the week. The periodic component is obtained from the weighted standard deviation estimator of Boudt et al. (2011) but computed on the centered log returns $r^*_{i\Delta}$ rather than $r_{i\Delta}$ as in Laurent and Shi (2020). See Boudt et al. (2011) and Laurent and Shi (2020) for a detailed presentation of this estimator. In the simulation studies, for simplicity, we assume $\hat{f}^*_{i\Delta}$ to be the same across days and set the estimated intraday periodicity to the true value plus zero mean noise. This is because for the estimation of intraday periodicity, one would need a long sample period. However, it is very computationally intensive to simulate long time spans at a one-second frequency. Therefore, the estimated intraday periodicity $\hat{f}^*_{i\Delta}$ is assumed to follow a $N(f^*_{i\Delta}, 0.01)$ distribution[8] and drawn from this distribution to account for estimation error of the periodicity on the jump test in our simulation study.

3.2.2 Feasible DF^J Test

The second step is to conduct the $DF^{\hat{J}}$ test based on the jump identification results. To do so, we replace K and $I^k_{i\Delta}$ in the regression model with their estimates such that

$$y_{i\Delta} = \alpha + \sum_{k=1}^{\hat{K}} \phi_k \hat{I}^k_{i\Delta} + \beta y_{(i-1)\Delta} + v_{i\Delta}. \tag{19}$$

The feasible test statistic $DF^{\hat{J}}$ is

$$DF^{\hat{J}} = \left(\ddot{\beta} - 1\right) \left[\frac{T \sum_{j=1}^{T} y^2_{j\Delta} - \left(\sum_{j=1}^{T} y_{(j-1)\Delta}\right)^2}{\sum_{j=1}^{T} \left(y_{i\Delta} - \ddot{\beta} y_{(i-1)\Delta} - \ddot{\alpha} - \sum_{k=1}^{\hat{K}} \ddot{\phi}_k \hat{I}^k_{i\Delta}\right)^2} \right]^{1/2},$$

where $\ddot{\alpha}$, $\ddot{\beta}$, and $\ddot{\phi}_k$ are the estimated OLS coefficients from regression (19).

Theorem 3.5 *The feasible unit root test $DF^{\hat{J}}$ has the following limiting properties. For a fixed time span N, as the sampling interval $\Delta \to 0$,*

$$DF^{\hat{J}} \Longrightarrow \frac{-\tilde{\Psi}_2 w_1 + \tilde{\Psi}_4}{\left(\tilde{\Psi}_3 - \tilde{\Psi}_2^2\right)^{1/2}} \equiv \Upsilon_2 \tag{20}$$

under the null hypothesis (9), and

$$DF^{\hat{J}} \Longrightarrow \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1}{\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2}} + c\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2} \equiv \Upsilon_2^A \tag{21}$$

under the alternative hypothesis (10).

The limiting distributions of $DF^{\hat{J}}$ are identical to those of DF^J, which demonstrates the asymptotic equivalence of the feasible and infeasible versions of the test. The proof of Theorem 3.2.2 is provided in Appendix D.

Remark 3.3 *In finite samples, the number of jumps might be over- or underestimated. In the case of the overidentification of jumps (i.e., $\hat{K} > K$), there are $(\hat{K} - K)$ redundant jump dummies in regression* (19)*. Those dummies will not have any asymptotic impact on the unit root test. The consistency of $\hat{\phi}_k$ remains, following directly from the proof of Theorem 3.5 (omitted here for brevity). The underidentification of jumps (i.e., $\hat{K} < K$) is most likely to happen when jump sizes are small, as the power of the jump test is higher when the jump size is larger (Laurent and Shi, 2020). The asymptotic impact of underidentification on the DF^J statistic is similar to that of jumps on the DF statistic (Theorem 3.3). However, the impact is expected to be marginal because the neglected jumps are small.*

[8]For the variance of $\hat{f}_{i\Delta}$ (i.e., the setting of 0.01), we first estimated the intraday periodicity with the 10-minute Nasdaq log prices in 1996 using the parametric approach proposed by Andersen and Bollerslev (1998b). Using the estimated parameters, we simulate data and run a Monte Carlo study to obtain the mean squared error of the fitted intraday periodicity.

Remark 3.4 *Using similar arguments as for the DF^J test, it is straightforward to show that the feasible version of the $DF(J)$ statistic, denoted by $DF(\hat{J})$, follows asymptotically the same distribution as the infeasible $DF(J)$ statistic.*

4 Simulation Studies

This section investigates the finite sample performance of the unit root test. We first consider a simple DGP with constant volatility (with or without jumps). We then extend this model with GARCH effects and intraday periodicity in the volatility and microstructure noise.

4.1 Constant Volatility

The data generating process is (9) under the null and (10) with $\theta \neq 0$ (or $\alpha \neq 1$) under the alternative. The value of θ is set to $\{-0.006, -0.004, -0.002\}$ under the mean reversion alternative and $\{0.002, 0.004, 0.006\}$ in the case of explosiveness. We consider the same parameter values for y_0, N, θ, and σ as in Section 2.3 and $\varepsilon_{i\Delta} \overset{i.i.d}{\sim} N(0,1)$. In the case of jumps, the locations of jumps $\tau_{i\Delta}^k$ are drawn randomly from a uniform distribution. The number of jumps is set to one per week. The magnitude of jumps ϕ_k is set to $\kappa\sigma$ with $\kappa = 2$ for positive jumps and $\kappa = -2$ for negative jumps.

Each day of simulated log prices consists of 23,400 observations, corresponding to one-second data over 6.5 hours. The one-second log prices are then aggregated to obtain data at the 10-minute frequency.[9] The nominal size of the tests is 5%, and the number of replications is $1,000$. We compare the performance of the new test statistics $DF^{\hat{J}}$ and $DF(\hat{J})$ with the conventional unit root test DF. Jump dummies are constructed by employing the test statistic of Laurent and Shi (2020) as presented in Section 3.2.1.

The empirical sizes and powers of both the left-sided (i.e., $H_1 : \theta < 0$) and the right-sided (i.e., $H_1 : \theta > 0$) tests are reported in Table 1. The left panel is for the test against the mean reversion alternative, while the right panel is for the test against explosiveness. The top panel reports the unit root test results in the absence of jumps in log prices. The performance of the three tests is almost identical in this setting, which is as expected and reassuring of good performance (low false identification rate) for the jump detection procedure. The empirical sizes of both tests are close to the nominal size. The power of each test increases as the process deviates further from the random walk. Note that when $N = 60$ and $\theta = -0.002$, the powers of the three left-sided tests are lower than the nominal size. The powers of the left-sided tests, however, increase as θ deviates further from zero (in the negative direction). This result is consistent with our observation from Figures 1, 3, and 5. That is, as θ decreases from 0 to negative values, the distribution of the test statistic first moves to the right of the null distribution before turning to the left. Moreover, the power of the tests increases with the time span N. Notably, when $\theta = -0.002$, the power of the left-sided tests increases from 0.6% to approximately 25% as the time span extends from one quarter ($N = 60$) to approximately one year ($N = 200$).

The middle and bottom panels are for the cases with positive and negative jumps, respectively. One can see that in the presence of jumps (either positive or negative), the left-sided DF test is undersized, while the right-sided DF test is severely oversized, which suggests a probability of falsely identifying jumps as explosive processes. The model under the null hypothesis (9) is

[9] The simulation results are qualitatively the same for 5-minute and 30-minute data and are therefore not reported to save space.

Table 1: Empirical sizes and powers of the unit root tests at the 10-minute frequency: constant volatility. The nominal size is 5%.

	$H_1 : \alpha < 1$				$H_1 : \alpha > 1$			
θ	-0.006	-0.004	-0.002	0	0	0.002	0.004	0.006
				No jumps				
$N = 60$								
DF	60.4	7.7	0.6	6.1	4.1	61.9	97.4	100.0
$DF^{\hat{J}}$	60.3	7.7	0.6	6.1	4.0	61.8	97.4	100.0
$DF(J)$	60.4	7.7	0.6	6.1	4.1	61.9	97.4	100.0
$N = 100$								
DF	100.0	53.8	3.5	5.4	5.2	83.6	100.0	100.0
$DF^{\hat{J}}$	100.0	53.8	3.5	5.4	5.2	83.6	100.0	100.0
$DF(J)$	100.0	53.8	3.5	5.4	5.1	83.6	100.0	100.0
$N = 200$								
DF	100.0	100.0	24.6	5.7	5.4	99.9	100.0	100.0
$DF^{\hat{J}}$	100.0	100.0	24.7	5.7	5.4	99.9	100.0	100.0
$DF(J)$	100.0	100.0	24.6	5.7	5.4	99.9	100.0	100.0
				Positive jumps				
$N = 60$								
DF	23.5	3.7	0.5	1.5	20.6	61.7	94.2	100.0
$DF^{\hat{J}}$	74.6	18.1	3.0	4.6	4.1	26.8	83.6	100.0
$DF(\hat{J})$	53.6	23.7	12.3	4.9	3.5	40.8	92.8	100.0
$N = 100$								
DF	89.3	18.8	1.6	1.2	28.6	81.1	100.0	100.0
$DF^{\hat{J}}$	100.0	69.1	9.0	5.3	4.3	48.5	100.0	100.0
$DF(\hat{J})$	92.9	54.2	22.1	5.7	4.5	59.4	100.0	100.0
$N = 200$								
DF	100.0	97.1	6.7	0.6	33.9	100.0	100.0	100.0
$DF^{\hat{J}}$	100.0	100.0	37.7	4.7	5.2	99.3	100.0	100.0
$DF(\hat{J})$	100.0	97.8	39.2	5.3	5.4	99.7	100.0	100.0
				Negative jumps				
$N = 60$								
DF	37.6	6.7	0.8	1.5	22.9	50.5	87.3	99.9
$DF^{\hat{J}}$	87.7	29.6	4.0	3.8	5.4	16.7	68.3	99.9
$DF(\hat{J})$	74.4	17.5	2.0	3.2	5.6	37.3	68.4	95.7
$N = 100$								
DF	98.2	36.7	2.0	1.1	30.0	66.0	99.9	100.0
$DF^{\hat{J}}$	100.0	89.4	16.1	4.1	5.9	27.9	99.2	100.0
$DF(\hat{J})$	100.0	82.2	10.8	4.8	5.4	44.6	91.5	100.0
$N = 200$								
DF	100.0	100.0	22.3	1.0	39.7	95.2	100.0	100.0
$DF^{\hat{J}}$	100.0	100.0	74.8	5.4	5.2	80.3	100.0	100.0
$DF(\hat{J})$	100.0	100.0	68.1	5.2	5.4	67.4	100.0	100.0

equivalent to a random walk process with breaks in the drift (due to the presence of jumps).[10] The results of the left-sided test echo the literature documenting observational equivalences between unit roots and structural breaks (e.g., Perron, 1990; Banerjee et al., 1992; Perron, 1997; Lumsdaine and Papell, 1997). Our result for the right-sided test is consistent with the findings of Phillips and Shi (2019), where a random drift martingale process is considered. Although one could visually separate negative jumps from an upward expanding explosive process (Phillips and Shi, 2019), there is no solution in the literature for distinguishing positive jumps from upward explosive processes.

Here, by including jump dummies in the model specifications, the $DF^{\hat{J}}$ test is able to isolate the impact of jumps while detecting breaks in the autoregressive coefficient. Both the $DF^{\hat{J}}$ and $DF(\hat{J})$ tests have reasonable sizes in all configurations. The size distortion of the DF tests translates into a lack of power for the left-sided DF test and more rejections for the right-sided DF test than the right-sided $DF^{\hat{J}}$ and $DF(\hat{J})$ tests. Interestingly, despite the low (absolute) values of the θ parameters considered in the simulation, $DF^{\hat{J}}$ and $DF(\hat{J})$ have good power against both alternatives. As expected, the powers increase with $|\theta|$ and N. Between DF^{J} and $DF(J)$, none of the tests uniformly dominates the other in terms of sizes and powers.

Recall that Figure 6 shows that when θ is negative and close to zero, all three tests have a probability of making a false positive rejection against the explosive alternative in the limit. The problem is much more severe for the DF and $DF(J)$ tests than the DF^{J} test. To verify these arguments in finite samples, we report the rejection frequencies of the left (resp. right) sided test while the true process is explosive (resp. stationary) on the left (resp. right) part of Table 2. The simulation results are consistent with our theoretical results. The left-sided tests have a near-zero probability of rejecting the null when the process is explosive (with positive θ). For the right-sided test, when the true process is mildly reverting (i.e., θ is negative and close to zero) and N is small (i.e., 60 and 100), all three tests wrongly detect an explosive pattern with high probability (approximately 33% when $N = 60$ and 20% when $N = 100$) in the absence of jumps. In the presence of jumps, $DF^{\hat{J}}$ clearly outperforms DF and $DF(\hat{J})$ in that it rejects the null hypothesis in favor of the wrong hypothesis much less frequently. For example, when $N = 60$, $\theta = -0.002$, and with positive jumps, the false rejection frequency of the right-sided DF^{J} test is 6.5%, compared with 34.4% of DF and 27.9% of $DF(J)$.

4.2 Time-varying Volatility and Microstructure Noise

To study the impact of time-varying volatility and microstructure noise, we consider more general model specifications. Under the null hypothesis, efficient log prices are now generated as follows:

$$y_{i\Delta} = \alpha_0 + \sum_{k=1}^{K} \phi_k I_{i\Delta}^k + y_{(i-1)\Delta} + \lambda_{i\Delta}\varepsilon_{i\Delta}, \tag{22}$$

where the volatility of log returns consists of a deterministic term $f_{i\Delta}$ and a stochastic component $\sigma_{i\Delta}$ such that

$$\lambda_{i\Delta} = f_{i\Delta}\sigma_{i\Delta}\sqrt{\Delta}.$$

We assume that f varies within the day but for simplicity restrict it to be the same across the N days in this simulation. To simulate a realistic periodic factor, we follow Laurent and Shi (2020) and take the estimated periodicity obtained using the parametric method proposed by Andersen and

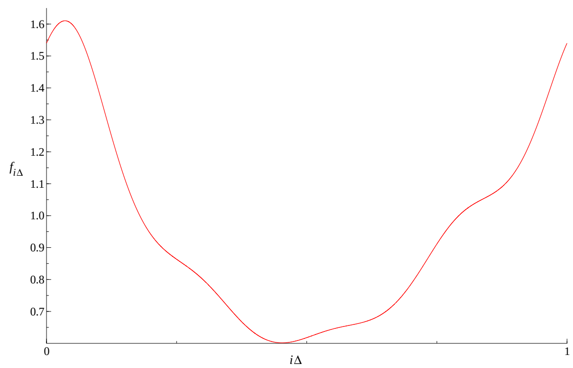
[10]This feature distinguishes jumps from bubbles, which are often modeled as a mildly explosive process (Phillips et al., 2011, 2015a; Phillips and Shi, 2018). The autoregressive coefficient of the mildly explosive process is greater than unity and takes the form of $\rho_T = 1 + cT^{-\alpha}$, with c being a constant, T being the sample size and $\alpha \in (0, 1)$.

Table 2: Rejection frequencies in favor of the wrong hypothesis at the 10-minute frequency: constant volatility. The nominal size is 5%.

	$H_1 : \alpha < 1$			$H_1 : \alpha > 1$		
	0.002	0.004	0.006	-0.006	-0.004	-0.002
			No jumps			
$N = 60$						
DF	0.1	0.0	0.0	0.2	6.9	32.8
$DF^{\hat{J}}$	0.1	0.0	0.0	0.2	6.9	32.9
$DF(\hat{J})$	0.1	0.0	0.0	0.2	6.9	32.9
$N = 100$						
DF	0.0	0.0	0.0	0.0	0.2	19.8
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.2	19.8
$DF(\hat{J})$	0.0	0.0	0.0	0.0	0.2	19.8
$N = 200$						
DF	0.0	0.0	0.0	0.0	0.0	1.3
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.0	1.3
$DF(\hat{J})$	0.0	0.0	0.0	0.0	0.0	1.3
			Positive jumps			
$N = 60$						
DF	0.0	0.0	0.0	1.9	15.9	34.4
$DF^{\hat{J}}$	0.2	0.0	0.0	0.0	0.6	6.5
$DF(\hat{J})$	0.2	0.0	0.0	3.8	15.9	27.9
$N = 100$						
DF	0.0	0.0	0.0	0.0	3.7	26.2
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.0	2.6
$DF(\hat{J})$	0.0	0.0	0.0	0.1	5.6	19.9
$N = 200$						
DF	0.0	0.0	0.0	0.0	0.0	8.9
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.0	0.0
$DF(\hat{J})$	0.0	0.0	0.0	0.0	0.0	7.3
			Negative jumps			
$N = 60$						
DF	0.2	0.0	0.0	0.8	10.5	31.8
$DF^{\hat{J}}$	0.8	0.0	0.0	0.0	0.2	5.0
$DF(\hat{J})$	8.5	1.4	0.0	0.0	0.9	10.9
$N = 100$						
DF	0.0	0.0	0.0	0.0	0.7	20.3
$DF^{\hat{J}}$	0.9	0.0	0.0	0.0	0.0	1.6
$DF(\hat{J})$	7.6	0.1	0.0	0.0	0.0	2.3
$N = 200$						
DF	0.0	0.0	0.0	0.0	0.0	2.4
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.0	0.0
$DF(\hat{J})$	1.5	0.0	0.0	0.0	0.0	0.0

Bollerslev (1998b) on the 10-minute Nasdaq stock price index from January 2, 1996, to December 8, 2017. The periodic component $f_{i\Delta}$ used in the simulations is plotted in Figure 7. It displays the usual diurnal pattern found in the volatility of intraday returns of most individual stocks and stock indices. The value of $f_{i\Delta}$ ranges from 0.6 to 1.6.

Figure 7: The simulated 1-second periodicity $f_{i\Delta}$ illustrated for one day.



The stochastic component is assumed to follow the GARCH(1,1) diffusion process of Nelson (1991), which has a discretized form of

$$\sigma^2_{i\Delta} = \delta_0 + \sigma^2_{(i-1)\Delta}(\beta_1 + \alpha_1 \varepsilon_{i\Delta}), \tag{23}$$

where $\delta_0 = \kappa\omega\Delta$, $\beta_1 = 1 - \kappa\Delta$, $\alpha_1 = \sqrt{2\lambda\kappa\Delta}$, and $\varepsilon_{i\Delta} \overset{i.i.d}{\sim} N(0,1)$. As in Andersen and Bollerslev (1998a), we choose the parameters $\kappa = 0.035$ and $\lambda = 0.296$ to simulate a log price process with realistic GARCH effects and set $\omega = 0.01^2$ such that $E(\sigma^2_{i\Delta}) = 0.01^2$ as in the previous simulations.

Under the alternative, we have

$$y_{i\Delta} = \alpha_0 + \sum_{k=1}^{K} \phi_k I^k_{i\Delta} + \beta_0 y_{(i-1)\Delta} + \lambda_{i\Delta}\varepsilon_{i\Delta}. \tag{24}$$

The settings of α_0, β_0 and jumps are the same as in the previous section.

Additionally, we assume that log prices are contaminated by microstructure noise such that $y^*_{i\Delta}$ is observed instead of the true efficient log price $y_{i\Delta}$, where

$$y^*_{i\Delta} = y_{i\Delta} + \varpi v_{i\Delta}$$

with $\varpi^2 = \xi\sqrt{\frac{1}{\Delta}\sum_{j=1}^{1/\Delta}\sigma^4_{i\Delta}}$ and $v_{i\Delta} \overset{i.i.d}{\sim} N(0,1)$. We set ξ to 0.0005.

Table 3 displays the rejection frequencies of the one-sided tests in the presence of microstructure noise, GARCH effects and intraday periodicity for a nominal size of 5%. The organization of the table is identical to that of Table 1. The results are qualitatively the same as in the case of constant volatility. In the absence of jumps, the three tests have similar sizes and powers. This finding suggests that heteroskedasticity and microstructure noise have little impact on the performance of the unit root tests when they are applied to 10-minute data and when $N \geq 60$.[11] As in the case of constant volatility, jumps have a strong impact on the performance of the DF test. Again, $DF^{\hat{J}}$

[11]Unreported simulation results suggest that time-varying volatility has a larger impact on test performance when the window size is small, e.g., when $N < 5$. In this case, as in Boswijk and Zu (2018), an adaptive wild bootstrap version of our test can be employed. We leave this extension for further research.

Table 3: Empirical sizes and powers of the unit root tests at the 10-minute frequency: GARCH effects, intraday periodicity and microstructure noise. The nominal size is 5%.

	$H_1: \alpha < 1$				$H_1: \alpha > 1$			
θ	-0.006	-0.004	-0.002	0	0	0.002	0.004	0.006
				No jumps				
$N = 60$								
DF	57.0	8.1	0.9	7.3	5.1	64.1	96.0	100.0
$DF^{\hat{J}}$	57.3	8.1	0.9	7.4	5.2	64.1	96.0	100.0
$DF(\hat{J})$	57.4	8.1	0.9	7.4	5.2	64.1	96.0	100.0
$N = 100$								
DF	98.9	48.2	2.4	6.6	5.3	82.1	100.0	100.0
$DF^{\hat{J}}$	98.9	48.0	2.4	6.7	5.2	82.1	100.0	100.0
$DF(\hat{J})$	98.9	48.0	2.4	6.7	5.3	82.2	100.0	100.0
$N = 200$								
DF	100.0	99.6	19.6	6.0	5.6	99.6	100.0	100.0
$DF^{\hat{J}}$	100.0	99.7	19.3	5.7	5.5	99.6	100.0	100.0
$DF(\hat{J})$	100.0	99.7	19.5	5.7	5.5	99.6	100.0	100.0
				Positive jumps				
$N = 60$								
DF	24.3	3.5	0.7	1.7	22.8	63.0	93.0	100.0
$DF^{\hat{J}}$	67.9	16.3	3.6	5.3	5.4	29.4	81.6	99.9
$DF(\hat{J})$	50.7	24.2	14.6	5.0	5.6	43.9	90.1	100.0
$N = 100$								
DF	24.3	3.5	0.7	1.7	22.8	63.0	93.0	100.0
$DF^{\hat{J}}$	67.9	16.3	3.6	5.3	5.4	29.4	81.6	99.9
$DF(\hat{J})$	50.7	24.2	14.6	5.0	5.6	43.9	90.1	100.0
$N = 200$								
DF	100.0	88.3	6.4	0.9	35.9	99.7	100.0	100.0
$DF^{\hat{J}}$	100.0	99.6	30.9	4.8	5.5	97.8	100.0	100.0
$DF(\hat{J})$	100.0	94.1	39.7	4.9	5.6	98.3	100.0	100.0
				Negative jumps				
$N = 60$								
DF	38.7	6.1	0.8	1.4	23.2	52.0	85.4	99.6
$DF^{\hat{J}}$	80.4	27.7	4.9	5.2	7.1	17.7	67.6	98.9
$DF(\hat{J})$	68.2	16.3	2.5	5.2	6.6	38.9	67.6	93.2
$N = 100$								
DF	93.3	32.8	2.6	1.9	30.2	63.2	98.8	100.0
$DF^{\hat{J}}$	99.9	84.3	13.4	5.4	4.1	25.9	98.2	100.0
$DF(\hat{J})$	99.7	74.0	8.9	4.9	4.5	42.2	88.4	99.9
$N = 200$								
DF	100.0	97.3	15.8	1.0	37.6	89.7	100.0	100.0
$DF^{\hat{J}}$	100.0	100.0	65.0	5.1	5.7	71.4	100.0	100.0
$DF(\hat{J})$	100.0	100.0	59.9	5.2	5.9	64.2	100.0	100.0

Table 4: Rejection frequencies in favor of the wrong hypothesis at the 10-minute frequency: GARCH effects, intraday periodicity and microstructure noise. The nominal size is 5%.

	$H_1 : \alpha < 1$			$H_1 : \alpha > 1$		
	0.002	0.004	0.006	-0.006	-0.004	-0.002
			No jumps			
$N = 60$						
DF	0.1	0.0	0.0	0.4	7.8	31.6
$DF^{\hat{J}}$	0.1	0.0	0.0	0.4	7.7	31.5
$DF(\hat{J})$	0.1	0.0	0.0	0.4	7.8	31.5
$N = 100$						
DF	0.0	0.0	0.0	0.0	0.4	20.2
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.4	20.5
$DF(\hat{J})$	0.0	0.0	0.0	0.0	0.4	20.4
$N = 200$						
DF	0.0	0.0	0.0	0.0	0.0	3.2
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.0	3.2
$DF(\hat{J})$	0.0	0.0	0.0	0.0	0.0	3.2
			Positive jumps			
$N = 60$						
DF	0.1	0.0	0.0	2.1	14.9	32.0
$DF^{\hat{J}}$	0.6	0.0	0.0	0.0	1.2	7.0
$DF(\hat{J})$	0.2	0.0	0.0	4.1	16.2	26.3
$N = 100$						
DF	0.0	0.0	0.0	0.0	3.4	26.8
$DF^{\hat{J}}$	0.1	0.0	0.0	0.0	0.0	2.8
$DF(\hat{J})$	0.0	0.0	0.0	0.0	4.9	19.6
$N = 200$						
DF	0.0	0.0	0.0	0.0	0.0	12.7
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.0	0.2
$DF(\hat{J})$	0.0	0.0	0.0	0.0	0.2	9.0
			Negative jumps			
$N = 60$						
DF	0.0	0.0	0.0	0.9	12.2	32.9
$DF^{\hat{J}}$	1.1	0.0	0.0	0.0	0.8	7.3
$DF(\hat{J})$	8.8	1.8	0.0	0.0	1.6	12.3
$N = 100$						
DF	0.0	0.0	0.0	0.1	1.6	21.4
$DF^{\hat{J}}$	0.3	0.0	0.0	0.0	0.0	1.0
$DF(\hat{J})$	9.1	0.4	0.1	0.0	0.0	2.1
$N = 200$						
DF	0.0	0.0	0.0	0.0	0.0	5.2
$DF^{\hat{J}}$	0.0	0.0	0.0	0.0	0.0	0.0
$DF(\hat{J})$	4.2	0.0	0.0	0.0	0.0	0.0

and $DF(\hat{J})$ have good power against both alternatives. There are cases where $DF^{\hat{J}}$ outperforms $DF(\hat{J})$ and vice versa.

Table 4 corresponds to Table 2 but with time-varying volatilities. As in Table 2, we see that in the presence of jumps, $DF^{\hat{J}}$ clearly outperforms $DF(\hat{J})$ in that the DF^{J} test rejects the null hypothesis in favor of the wrong hypothesis much less frequently. For this reason, and since there is tremendous evidence of the presence of jumps in high-frequency asset prices, we therefore recommend the use of $DF^{\hat{J}}$ in empirical applications.

5 Empirical Application

The primary purpose of this section is to show the impact of jumps on unit root testing on intraday data by comparing the performance of the standard DF test and the proposed $DF^{\hat{J}}$ test. The $DF(\hat{J})$ test is not considered because it has a much higher probability of rejecting against wrong hypotheses (as discussed). We investigate the dynamics of the 10-minute log prices of the Nasdaq composite index over two sample periods (1999-05-01 to 2000-06-30 and 2015-05-01 to 2016-01-31). The data are downloaded from Thomson Reuters DataScope and displayed in Figure 8.

The first period falls in the famous dot-com bubble period (Phillips et al., 2011; Shi and Song, 2016). It has been widely recognized that asset prices exhibit explosive dynamics in the presence of speculative bubbles (Diba and Grossman 1988; Phillips et al. 2011; Phillips et al. 2015a,b). Evidence of speculative bubbles has been detected in various markets with low-frequency data (daily, weekly or monthly).[12] The dot-com bubble is the most prominent episode. We can see a dramatic increase in the Nasdaq stock price in the second half of 1999. The market peaked on March 10, 2000, followed by a sharp downturn.

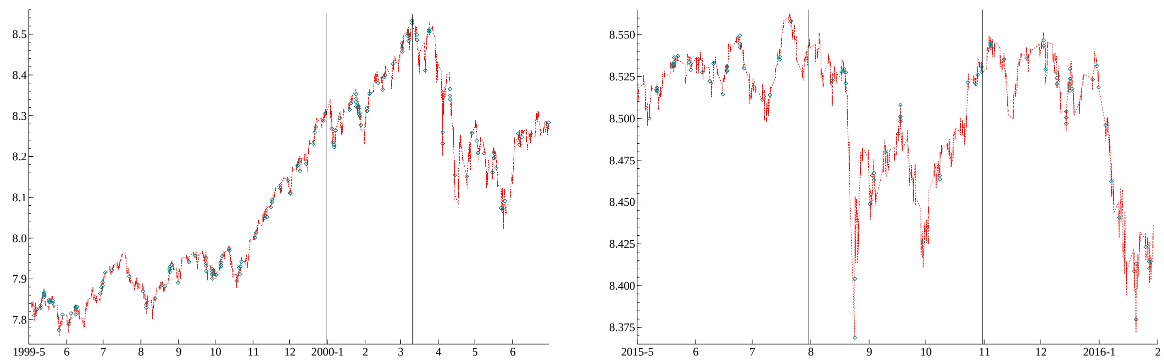
The second episode is around the 2015-2016 stock market sell-off, triggered by the bursting of the Chinese stock market bubble on 12 June 2015. We observe dramatic turbulence in the Nasdaq stock market between August and October. In particular, the Nasdaq market dropped 15.59% from 19 August to 24 August. It recovered to approximately the same level as before the crash by the end of October 2015. Various models have been proposed for capturing the dynamics of a bubble bursting and market crashes (e.g., a mildly stationary process of Phillips and Shi, 2018 and the random drift martingale process of Phillips and Shi, 2019) but yet to be tested empirically.

We first apply the jump test of Laurent and Shi (2020) (with the same critical values as in the simulation studies) to the two data series to create the jump dummies. In empirical applications on jump detection with high-frequency data, overnight returns are often removed because they convey information on a more extended period than the other returns (i.e., 17.5 hours rather than 10 minutes in our case). Removing the first observation of each day is obviously not a solution for unit root tests with log prices. Therefore, we keep the first observations of the days (and hence overnight returns) for both the jump and unit root tests. The variance of overnight returns can be captured by the periodicity component, which is taken care of in our test for jumps. Moreover, we show in our simulations that the periodicity of variances does not have a significant impact on unit root tests when N is larger than 60. Therefore, the inclusion of overnight returns is not expected to affect our test outcomes. The identified jumps are marked with diamonds in Figure 8. There are 149 jumps during the first sample (i.e., 1999-05-1 and 2000-06-30) and 91 during the second one (i.e., 2015-05-01 and 2016-01-31). It is clear from this figure that jumps are not rare events. Furthermore, some of the detected jumps are very large and therefore expected to have an impact on the DF test.

[12]See, for example, Phillips et al. (2011), Gutierrez (2012); Fantazzini (2016); Etienne et al. (2013); Pavlidis et al. (2016); Hu and Oxley (2018); Shi (2017).

Figure 8: The 10-minute log prices of the Nasdaq composite index for two sample periods. The diamonds indicate the jumps used in the $DF^{\hat{J}}$ test. The vertical lines indicate the cutoff dates of each subsample.

(a) 1999-05-01 to 2000-06-30 (b) 2015-05-01 to 2016-01-31



For the unit root tests, we divide each period into three subsamples, guided by the important events discussed above. The cutoff dates of the subsamples are marked by vertical lines in Figure 8. The unit root test statistics DF and $DF^{\hat{J}}$ and their corresponding asymptotic critical values for the left- and right-sided tests (5% and 95%) are reported in Table 5, along with the exact dates of each subsample and the number of jumps detected in each subsample.[13]

For the dot-com bubble period, both procedures detect explosive dynamics in the Nasdaq stock market in the first subsample (between May 1999 and December 1999).[14] In the second subsample, while the null hypothesis of a random walk is still rejected against explosiveness with the DF test, the DF^{J} test does not reject the null against the explosive alternative. The distinct outcomes of the DF and $DF^{\hat{J}}$ tests could potentially be explained by our findings in Sections 3 and 4. That is, jumps could lead to spurious rejections of the DF test against the alternative hypothesis of explosiveness. In contrast, the DF^{J} test, which accounts for the presence of jumps, has satisfactory performance under this circumstance. The $DF^{\hat{J}}$ test suggests that the process returns to a random walk in the period from 2000-01-01 to 2000-03-10 before reaching the peak of the bubble episode. This result has important implications for traders who have every intension to withdraw from the market before bubbles collapse. Interestingly, the two tests again agree when applied to the third subsample, spanning between 11 March 2000 and 30 June 2000. Indeed, we fail to reject the null hypothesis during this period using both tests, suggesting that the bursting of the dot-com bubble follows a random walk pattern.

Similarly, for the second sample period (2015-2016), DF and $DF^{\hat{J}}$ yield consistent results in the first and third subsamples but draw different conclusions in the second subsample. Both tests

[13]Since the jump dummies are orthogonal to each other and $\hat{\phi}_k$ is asymptotically normally distributed (according to Theorem 3.2), standard t-tests can be used to eliminate insignificant jump dummies in Equation (10). As the results are qualitatively the same when including all jump dummies or only dummies that are significant at the conventional significance levels, we only report the results with all jump dummies. Unreported simulations reveal similar sizes and powers for the $DF^{\hat{J}}$ test when the regression includes all jump dummies (including redundant ones) or only significant ones.

[14]The presence of explosive dynamics in asset prices does not imply the existence of bubbles. An additional necessary step is to control for the impact of market fundamentals. Stock market fundamentals are often proxied by dividends or earnings, which are unfortunately not available at such a high frequency. Therefore, we do not refer to the explosive dynamics as bubbles in this paper.

Table 5: Results of the DF and $DF^{\hat{J}}$ tests

Start	End	DF	left cv	right cv	$DF^{\hat{J}}$	left cv	right cv	Jumps
1999-05-01	1999-12-31	**1.639**	-2.857	-0.089	**2.215**	-2.819	0.062	84
2000-01-01	2000-03-10	**0.160**	-2.857	-0.089	0.312	-2.672	0.404	34
2000-03-11	2000-06-30	-2.008	-2.857	-0.089	-1.937	-2.809	0.005	31
2015-05-01	2015-07-31	-2.357	-2.857	-0.089	-2.426	-2.823	-0.038	32
2015-08-01	2015-10-15	-2.406	-2.857	-0.089	**-3.026**	-2.714	0.268	22
2015-10-16	2016-01-31	-0.788	-2.857	-0.089	-0.120	-2.746	0.057	37

Note: The statistics highlighted in bold are significant at the 5% level using either the left- or right-sided critical values. The figures in the columns left cv and right cv are the critical values used for the corresponding left-sided and right-sided tests, respectively. The figures in the last column correspond to the number of jumps detected using the test of Laurent and Shi (2020).

conclude that the log Nasdaq price follows a unit root process in the first and third subsamples. For the turbulent 2.5 months in 2015 (from 1 August to 15 October), the $DF^{\hat{J}}$ rejects the null against the mean reversion alternative, while the DF test fails to do so. Again, this finding is not surprising, as we find in Sections 3 and 4 that jumps decrease the power of the DF test against mean reversion, and we observe jumps with extremely large magnitudes over this period (e.g., at the opening of 24 August 2015). Interestingly, unlike the bursting of the dot-com bubble, the $DF^{\hat{J}}$ test suggests that the stock market crash in late 2015 follows a mean reversion process.

Finally, although the availability of high-frequency data allows us to conduct the unit root tests using data over short time periods and hence reduce the probability of having structural breaks within the sample period, it does not completely rule out this possibility. The performance of the unit root tests in the presence of structural breaks remains unknown. One could potentially account for structural breaks with the model proposed by Jiang et al. (2017) but extended to allow for jumps. Here, we provide examples of cases when both DF and $DF^{\hat{J}}$ reject the unit root null hypothesis in favor of the same alternative and, more important, when they contradict, with carefully divided subsamples. A comprehensive analysis of the structural break issue and a more extensive empirical application over a longer sample period is left for future work.

6 Conclusion

This paper provides an efficient tool for detecting deviations of asset prices from a random walk with intraday high-frequency data. The proposed tool is based on unit root tests but takes the empirical features of high-frequency data (particularly jumps) into consideration. The null hypothesis is a random walk, while the alternative hypothesis is either mean reversion or explosiveness. With the in-fill asymptotic, we show that the conventional DF tests could lead to severe size distortions in the presence of jumps, according to both the asymptotic and simulation results.

We propose two new tests that account for the possible presence of jumps, denoted DF^J and $DF(J)$ for the unfeasible version and $DF^{\hat{J}}$ and $DF(\hat{J})$ for the feasible one. The limiting distributions of the new test statistics under both the null and the alternative are provided. Both tests depend on nuisance parameters for which we propose consistent estimators. Importantly, we show the asymptotic equivalence between the infeasible (i.e., assuming true jump occurrences) and feasible (i.e., relying on a test to identify jumps) versions of the tests. Simulation results reveal the satisfactory performance (in terms of size and power) of the new tests but also that $DF(\hat{J})$ tends to reject too often against the wrong alternative when the process is mildly mean-reverting ($\theta < 0$ but close to zero), and N is relatively small. Therefore, we recommend the use of $DF^{\hat{J}}$ for empirical applications.

Furthermore, we show via simulations that conditional heteroskedasticity, intraday periodicity, and microstructure noise do not affect the finite sample performance of the tests when the test window is applied to one quarter of data (or more), and the sampling frequency is 10 minutes or lower.

We apply the conventional DF test and $DF^{\hat{J}}$ to the 10-minute log prices of the Nasdaq composite index around the peak of the dot-com bubble (1999-2000) and the 2015-2016 stock market sell-off periods. Both tests reject the null against the explosive alternative in late 1999. The two unit root tests, however, provide contradictory results in the early 2000s before the bursting of the dot-com bubble and in late 2015 when the market experienced turbulence. We attribute these differences to the lack of power of the left-sided DF test and the oversize issue of the right-sided DF test when jumps are ignored. The $DF^{\hat{J}}$ test suggests that log Nasdaq prices switch back to a random walk dynamic (from being explosive) as the peak of the bubble approaches. Additionally, the dot-com bubble collapses in a random walk fashion, while the 2015 stock market crash follows a mean-reverting process.

References

Ahn, C.M., Thompson, H.E., 1988. Jump-diffusion processes and the term structure of interest rates. Journal of Finance 43, 155–174.

Aït-Sahalia, Y., Mykland, P., Zhang, L., 2005. How often to sample a continuous-time process in the presence of market microstructure noise. Review of Financial Studies 18, 351.

Ait-Sahalia, Y., Yu, J., 2009. High frequency market microstructure noise estimates and liquidity measures. Annals of Applied Statistics 3, 422–457.

Andersen, T., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets. Journal of Empirical Finance 4, 115–158.

Andersen, T., Bollerslev, T., 1998a. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. International Economic Review 39, 885–905.

Andersen, T., Bollerslev, T., Diebold, F., 2007a. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. The Review of Economics and Statistics 89, 701–720.

Andersen, T.G., Bollerslev, T., 1998b. Deutsche mark-dollar volatility: intra-day activity patterns, macroeconomic announcements and longer run dependencies. Journal of Finance 53, 219–265.

Andersen, T.G., Bollerslev, T., 1998b. Deutsch mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. Journal of Finance 53, 219–265.

Andersen, T.G., Bollerslev, T., Dobrev, D., 2007b. No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and iid noise: Theory and testable distributional implications. Journal of Econometrics 138, 125–180.

Bajgrowicz, P., Scaillet, O., Treccani, A., 2015. Jumps in high-frequency data: Spurious detections, dynamics, and news. Management Science 62, 2198–2217.

Balvers, R., Wu, Y., Gilliland, E., 2000. Mean reversion across national stock markets and parametric contrarian investment strategies. The Journal of Finance 55, 745–772.

Banerjee, A., Lumsdaine, R.L., Stock, J.H., 1992. Recursive and sequential tests of the unit-root and trend-break hypotheses: theory and international evidence. Journal of Business & Economic Statistics 10, 271–287.

Bauwens, L., Hafner, C., Laurent, S., 2012. Handbook of volatility models and their applications. volume 3. John Wiley & Sons.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307–327.

Boswijk, H.P., Zu, Y., 2018. Adaptive wild bootstrap tests for a unit root with nonstationary volatility. The Econometrics Journal 21, 87–113.

Boudt, K., Croux, C., Laurent, S., 2011. Robust estimation of intraweek periodicity in volatility and jump detection. Journal of Empirical Finance 18, 353–367.

Brooks, C., Katsaris, A., 2005. A three-regime model of speculative behaviour: Modelling the evolution of the S&P 500 composite index. The Economic Journal 115, 767–797.

Chaudhuri, K., Wu, Y., 2003. Random walk versus breaking trend in stock prices: Evidence from emerging markets. Journal of Banking & Finance 27, 575–592.

Clemente, J., Montañés, A., Reyes, M., 1998. Testing for a unit root in variables with a double change in the mean. Economics Letters 59, 175–182.

Diba, B.T., Grossman, H.I., 1988. Explosive rational bubbles in stock prices? The American Economic Review 78, 520–530.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74, 427–431.

Dickey, D.A., Fuller, W.A., 1981. Likelihood ratio statistics for autoregressive time series with a unit root. Econometrica 49, 1057–1072.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of united kingdom inflation. Econometrica 50, 987–1007.

Etienne, X.L., Irwin, S.H., Garcia, P., 2013. Bubbles in food commodity markets: Four decades of evidence. Journal of International Money and Finance 97, 65–87.

Evans, G.W., 1991. Pitfalls in testing for explosive bubbles in asset prices. The American Economic Review 81, 922–930.

Fama, E.F., French, K.R., 1988. Permanent and temporary components of stock prices. Journal of Political Economy 96, 246–273.

Fantazzini, D., 2016. The oil price crash in 2014/15: Was there a (negative) financial bubble? Energy Policy 96, 383–396.

Gatev, E., Goetzmann, W.N., Rouwenhorst, K.G., 2006. Pairs trading: Performance of a relative-value arbitrage rule. Review of Financial Studies 19, 797–827.

Guenster, N., Kole, E., 2013. Bubbles and long-term investors.

Gutierrez, L., 2012. Speculative bubbles in agricultural commodity markets. European Review of Agricultural Economics 40, 217–238.

Hamilton, J.D., 1994. Time series analysis. volume 2. Princeton university press Princeton.

Harvey, D., Leybourne, S., Zu, Y., 2019. Testing explosive bubbles with time-varying volatility. Econometric Review 38, 1131–1151.

Homm, U., Breitung, J., 2012. Testing for speculative bubbles in stock markets: a comparison of alternative methods. Journal of Financial Econometrics 10, 198–231.

Hu, Y., Oxley, L., 2018. Do 18th century bubbles survive the scrutiny of 21st century time series econometrics? Economics Letters 162, 131–134.

Jiang, L., Wang, X., Yu, J., 2017. In-fill asymptotic theory for structural break point in autoregression: A unified theory.

Jiang, L., Wang, X., Yu, J., 2018. New distribution theory for the estimation of structural break point in mean. Journal of Econometrics 205, 156–176.

Kim, J., Park, J.Y., 2019. Unit root, mean reversion and nonstationarity in financial time series. Working paper .

Kim, M.J., Nelson, C.R., Startz, R., 1991. Mean reversion in stock prices? A reappraisal of the empirical evidence. The Review of Economic Studies 58, 515–528.

Kou, S.G., 2002. A jump-diffusion model for option pricing. Management Science 48, 1086–1101.

Kwiatkowski, D., Phillips, P.C., Schmidt, P., Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? Journal of Econometrics 54, 159–178.

Laurent, S., Shi, S., 2020. Volatility Estimation and Jump Detection for Drift-diffusion Processes. Technical Report. Available at: https://doi.org/10.1016/j.jeconom.2019.12.004. Forthcoming in Journal of Econometrics.

Lee, J., Strazicich, M.C., 2001. Break point estimation and spurious rejections with endogenous unit root tests. Oxford Bulletin of Economics and Statistics 63, 535–558.

Lee, J., Strazicich, M.C., 2003. Minimum lagrange multiplier unit root test with two structural breaks. Review of Economics and Statistics 85, 1082–1089.

Lee, S.S., 2011. Jumps and information flow in financial markets. The Review of Financial Studies 25, 439–479.

Lee, S.S., 2012. Jumps and information flow in financial markets. Review of Financial Studies 15 (2), 439–479.

Lee, S.S., Mykland, P.A., 2008. Jumps in financial markets: A new nonparametric test and jump dynamics. Review of Financial Studies 21, 2535–2563. doi:http://dx.doi.org/10.1093/rfs/hhm056.

Lo, A.W., MacKinlay, A.C., 1988. Stock market prices do not follow random walks: Evidence from a simple specification test. Review of Financial Studies 1, 41–66.

Lumsdaine, R.L., Papell, D.H., 1997. Multiple trend breaks and the unit-root hypothesis. Review of Economics and Statistics 79, 212–218.

Mancini, C., 2011. Jumps, in: Bauwens, L., Hafner, C., Laurent, S. (Eds.), Wiley Handbook in Financial Engineering and Econometrics: Volatility Models and Their Applications, Wiley, New York.

McQueen, G., 1992. Long-horizon mean-reverting stock prices revisited. Journal of Financial and Quantitative Analysis 27, 1–18.

Merton, R.C., 1976. Option pricing when underlying stock returns are discontinuous. Journal of Financial Economics 3, 125–144.

Merton, R.C., 1980. On estimating the expected return on the market: An exploratory investigation. Journal of Financial Economics 8, 323–361.

Miller, M.H., Muthuswamy, J., Whaley, R.E., 1994. Mean reversion of standard & poor's 500 index basis changes: Arbitrage-induced or statistical illusion? The Journal of Finance 49, 479–513.

Milunovich, G., Shi, S.P., Tan, D., 2017. Bubble detection and sector trading in real time. Available at SSRN: http://dx.doi.org/10.2139/ssrn.2827051 .

Narayan, P.K., Sharma, S.S., Phan, D.H.B., 2016. Asset price bubbles and economic welfare. International Review of Financial Analysis 44, 139–148.

Nelson, C.R., Plosser, C.R., 1982. Trends and random walks in macroeconmic time series: some evidence and implications. Journal of Monetary Economics 10, 139–162.

Nelson, D.B., 1991. ARCH models as diffusion approximations. Journal of Econometrics 45, 7–38.

Pavlidis, E., Yusupova, A., Paya, I., Peel, D., Martínez-García, E., Mack, A., Grossman, V., 2016. Episodes of exuberance in housing markets: In search of the smoking gun. The Journal of Real Estate Finance and Economics 53, 419–449.

Perron, P., 1989. The great crash, the oil price shock, and the unit root hypothesis. Econometrica 57, 1361–1401.

Perron, P., 1990. Testing for a unit root in a time series with a changing mean. Journal of Business & Economic Statistics 8, 153–162.

Perron, P., 1991. A continuous time approximation to the unstable first-order autoregressive process: the case without an intercept. Econometrica 59, 211–236.

Perron, P., 1997. Further evidence on breaking trend functions in macroeconomic variables. Journal of Econometrics 80, 355 –385.

Phillips, P., Yu, J., 2013. Bubble or roller coaster in world stock markets. The Business Times. June 28.

Phillips, P.C., 1987a. Time series regression with a unit root. Econometrica , 277–301.

Phillips, P.C., 1987b. Towards a unified asymptotic theory for autoregression. Biometrika , 535–547.

Phillips, P.C., Perron, P., 1988. Testing for a unit root in time series regression. Biometrika , 335 –346.

Phillips, P.C., Shi, S., 2018. Financial bubble implosion and reverse regression. Econometric Theory 34, 705–753.

Phillips, P.C., Shi, S., 2019. Detecting financial collapse and ballooning sovereign risk. Oxford Bulletin of Economics and Statistics 81, 1336–1361.

Phillips, P.C., Shi, S., Yu, J., 2015a. Testing for multiple bubbles: Historical episodes of exuberance and collapse in the S&P 500. International Economic Review 56, 1043–1078.

Phillips, P.C., Shi, S., Yu, J., 2015b. Testing for multiple bubbles: Limit theory of real time detectors. International Economic Review 56, 1079–1134.

Phillips, P.C., Wu, Y., Yu, J., 2011. Explosive behavior in the 1990s NASDAQ: When did exuberance escalate asset values? International Economic Review 52, 201–226.

Phillips, P.C., Yu, J., 2011. Dating the timeline of financial bubbles during the subprime crisis. Quantitative Economics 2, 455–491.

Poterba, J.M., Summers, L.H., 1988. Mean reversion in stock prices: Evidence and implications. Journal of Financial Economics 22, 27–59.

Richards, A.J., 1997. Winner-loser reversals in national stock market indices: can they be explained? The Journal of Finance 52, 2129–2144.

Richardson, M., 1993. Temporary components of stock prices: A skeptic's view. Journal of Business & Economic Statistics 11, 199–207.

Said, S.E., Dickey, D.A., 1984. Testing for unit roots in autoregressive-moving average models of unknown order. Biometrika 71, 599–607.

Schmidt, P., Phillips, P.C., 1992. Lm tests for a unit root in the presence of deterministic trends. Oxford Bulletin of Economics and Statistics 54, 257–287.

Serban, A.F., 2010. Combining mean reversion and momentum trading strategies in foreign exchange markets. Journal of Banking & Finance 34, 2720–2727.

Shi, S., 2017. Speculative bubbles or market fundamentals? an investigation of US regional housing markets. Economic Modelling 66, 101–111.

Shi, S., Song, Y., 2016. Identifying speculative bubbles using an infinite hidden Markov model. Journal of Financial Econometrics 14, 159–184.

Shiller, R.J., Perron, P., 1985. Testing the random walk hypothesis: Power versus frequency of observation. Economics Letters 18, 381–386.

Tao, Y., Phillips, P.C., Yu, J., 2019. Random coefficient continuous systems: Testing for extreme sample path behavior. Journal of Econometrics 209, 208–237.

Taylor, S., Xu, X., 1997. The incremental volatility information in one million foreign exchange quotations. Journal of Empirical Finance 4, 317–340.

Taylor, S.J., 1994. Modeling stochastic volatility: A review and comparative study. Mathematical Finance 4, 183–204.

Vogelsang, T.J., Perron, P., 1998. Additional tests for a unit root allowing for a break in the trend function at an unknown time. International Economic Review , 1073–1100.

Wang, X., Yu, J., 2016. Double asymptotics for explosive continuous time models. Journal of Econometrics 193, 35–53.

Yu, J., 2014. Econometric analysis of continuous time models: A survey of Peter Phillips's work and some new results. Econometric Theory 30, 737–774.

Zhou, Q., Yu, J., 2015. Asymptotic theory for linear diffusions under alternative sampling schemes. Economics Letters 128, 1–5.

Zivot, E., Andrews, D.W.K., 2002. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. Journal of Business & Economic Statistics 20, 25–44.

Appendix A: Proof of Remark 2.2

Proof. The DF statistic can be rewritten as

$$
\begin{aligned}
DF &= \left[\left(\hat{\beta}-\beta_0\right)+(\beta_0-1)\right]\left[\frac{T\sum_{i=1}^{T} y_{i\Delta}^2-\left(\sum_{i=1}^{T} y_{(i-1)\Delta}\right)^2}{\sum_{i=1}^{T}\left(y_{i\Delta}-\hat{\alpha}-\hat{\beta}y_{(i-1)\Delta}\right)^2}\right]^{1/2} \\
&= t_{\hat{\beta}}+(\beta_0-1)\left[\frac{T\sum_{i=1}^{T} y_{i\Delta}^2-\left(\sum_{i=1}^{T} y_{(i-1)\Delta}\right)^2}{\sum_{i=1}^{T}\left(y_{i\Delta}-\hat{\alpha}-\hat{\beta}y_{(i-1)\Delta}\right)^2}\right]^{1/2}.
\end{aligned}
$$

(1) When $\theta > 0$, from Wang and Yu (2016), under the DGP of (2) and the double asymptotic scheme,

$$
\begin{aligned}
t_{\hat{\beta}} &\Longrightarrow N(0,1), \\
\frac{e^{2\theta\Delta}-1}{e^{2\theta N}}\sum_{i=1}^{T} y_{i\Delta}^2 &\Longrightarrow (y_0-\mu+\sigma Z_1)^2, \\
\frac{e^{\theta\Delta}-1}{e^{\theta N}}\sum_{i=1}^{T} y_{i\Delta} &\Longrightarrow \sigma Z_1 + y_0-\mu, \\
\frac{1}{N}\sum_{i=1}^{T}\left(y_{i\Delta}-\hat{\alpha}-\hat{\beta}y_{(i-1)\Delta}\right)^2 &\to \sigma^2,
\end{aligned}
$$

with $Z_1 \sim N(0, \frac{1}{2\theta})$. Therefore,

$$(\beta_0 - 1)\frac{\left[T\sum_{i=1}^T y_{i\Delta}^2 - \left(\sum_{i=1}^T y_{(i-1)\Delta}\right)^2\right]^{1/2}}{\left[\sum_{i=1}^T \left(y_{i\Delta} - \hat{\alpha} - \hat{\beta} y_{(i-1)\Delta}\right)^2\right]^{1/2}} = \left(e^{\theta\Delta} - 1\right)\left[\frac{T\frac{e^{2\theta N}}{e^{2\theta\Delta}-1}\left(y_0 - \mu + \sigma Z_1\right)^2}{\sigma^2 N}\right]^{1/2}[1 + o_p(1)]$$

$$\sim e^{\theta N}\sqrt{\frac{\theta}{2}}\frac{1}{\sigma}\left|y_0 - \mu + \sigma Z_1\right| \to +\infty.$$

It follows that

$$DF = (\beta_0 - 1)\frac{\left[T\sum_{i=1}^T y_{i\Delta}^2 - \left(\sum_{i=1}^T y_{(i-1)\Delta}\right)^2\right]^{1/2}}{\left[\sum_{i=1}^T \left(y_{i\Delta} - \hat{\alpha} - \hat{\beta} y_{(i-1)\Delta}\right)^2\right]^{1/2}}[1 + o_p(1)] \sim e^{\theta N}\sqrt{\frac{\theta}{2}}\frac{1}{\sigma}\left|y_0 - \mu + \sigma Z_1\right| \to +\infty.$$

(2) Consider the case of $\theta < 0$. From Wang and Yu (2016), assuming $E\left|\varepsilon_{1\Delta}\right|^{2+\delta} < \infty$ for some $\delta > 0$, under the DGP of (2) and the double asymptotic scheme,

$$\begin{aligned} t_{\hat{\beta}} &\Longrightarrow N(0,1), \\ \frac{1}{T}\sum_{i=1}^T y_{i\Delta}^2 &\Longrightarrow -\frac{1}{2\theta} + \frac{\mu^2}{\sigma^2}, \\ \frac{1}{T}\sum_{i=1}^T y_{i\Delta} &\Longrightarrow \frac{\mu}{\sigma}, \\ \frac{1}{N}\sum_{i=1}^T \left(y_{i\Delta} - \hat{\alpha} - \hat{\beta} y_{(i-1)\Delta}\right)^2 &\to \sigma^2. \end{aligned}$$

Therefore,

$$(\beta_0 - 1)\frac{\left[T\sum_{i=1}^T y_{i\Delta}^2 - \left(\sum_{i=1}^T y_{(i-1)\Delta}\right)^2\right]^{1/2}}{\left[\sum_{i=1}^T \left(y_{i\Delta} - \hat{\alpha} - \hat{\beta} y_{(i-1)\Delta}\right)^2\right]^{1/2}} = \frac{T}{\sigma N^{1/2}}\left(e^{\theta\Delta} - 1\right)\left[\frac{1}{T}\sum_{i=1}^T y_{i\Delta}^2 - \left(\frac{1}{T}\sum_{i=1}^T y_{i\Delta}\right)^2\right]^{1/2}[1 + o_p(1)]$$

$$\sim N^{1/2}\frac{\theta}{\sigma}\left(-\frac{1}{2\theta}\right)^{1/2} \to -\infty.$$

It follows that when $\theta < 0$

$$DF = (\beta_0 - 1)\frac{\left[T\sum_{i=1}^T y_{i\Delta}^2 - \left(\sum_{i=1}^T y_{(i-1)\Delta}\right)^2\right]^{1/2}}{\left[\sum_{i=1}^T \left(y_{i\Delta} - \hat{\alpha} - \hat{\beta} y_{(i-1)\Delta}\right)^2\right]^{1/2}}[1 + o_p(1)] \sim N^{1/2}\frac{\theta}{\sigma}\left(-\frac{1}{2\theta}\right)^{1/2} \to -\infty.$$

■

Appendix B: Asymptotics of Models with Jumps

Proof of Lemma 3.1. The null hypothesis (9) can be rewritten as

$$y_{i\Delta} = \sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k} + \sigma\sqrt{\Delta}\sum_{j=1}^{i}\varepsilon_{j\Delta} + y_0. \tag{25}$$

(a) We have

$$y_{T\Delta} = \sum_{j=1}^{T}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k} + \sigma\sqrt{\Delta}\sum_{j=1}^{T}\varepsilon_{j\Delta} + y_0 \Longrightarrow \sigma N^{1/2}\left[\sum_{k=1}^{K}\frac{\phi_k}{\sigma N^{1/2}} + \Psi_1\right] \equiv \sigma N^{1/2}\tilde{\Psi}_1$$

since

$$\sum_{j=1}^{T}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k} = \sum_{k=1}^{K}\phi_k\sum_{j=1}^{T}I_{j\Delta}^{k} = \sum_{k=1}^{K}\phi_k$$

and from the proof Lemma 2.1 in the online supplement

$$\sigma\sqrt{\Delta}\sum_{j=1}^{T}\varepsilon_{j\Delta} + y_0 \Longrightarrow \sigma N^{1/2}\Psi_1.$$

(b) The quantity

$$\begin{aligned} T^{-1}\sum_{i=1}^{T}y_{i\Delta} &= T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k} + T^{-1}\sum_{i=1}^{T}\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i}\varepsilon_{j\Delta} + y_0\right) \\ &\Longrightarrow \sigma N^{1/2}\left[\sum_{k=1}^{K}\frac{\phi_k}{\sigma N^{1/2}}(1-r_k) + \Psi_2\right] \equiv \sigma N^{1/2}\tilde{\Psi}_2 \end{aligned}$$

since

$$\begin{aligned} T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k} &= T^{-1}\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}\sum_{j=1}^{i}I_{j\Delta}^{k} \\ &= T^{-1}\sum_{k=1}^{K}\phi_k\left[I_{1\Delta}^{k} + (I_{1\Delta}^{k}+I_{2\Delta}^{k}) + \cdots + (I_{1\Delta}^{k}+I_{2\Delta}^{k}+\cdots+I_{T\Delta}^{k})\right] \\ &= T^{-1}\sum_{k=1}^{K}\phi_k(T-\tau_k) = \sum_{k=1}^{K}\phi_k(1-r_k) \end{aligned}$$

and from the proof Lemma 2.1 in the online supplement

$$T^{-1}\sum_{i=1}^{T}\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i}\varepsilon_{j\Delta} + y_0\right) \Longrightarrow \sigma N^{1/2}\Psi_2.$$

(c) The quantity

$$T^{-1}\sum_{i=1}^{T}y_{i\Delta}^{2} = T^{-1}\sum_{i=1}^{T}\left[\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k} + \sigma\sqrt{\Delta}\sum_{j=1}^{i}\varepsilon_{j\Delta} + y_0\right]^2$$

$$= T^{-1}\sum_{i=1}^{T}\left(\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right)^2 + T^{-1}\sum_{i=1}^{T}\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i}\varepsilon_{j\Delta}+y_0\right)^2$$
$$+ 2T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i}\varepsilon_{j\Delta}+y_0\right).$$

The first term

$$\begin{aligned}
T^{-1}\sum_{i=1}^{T}\left(\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right)^2 &= T^{-1}\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k\sum_{j=1}^{i} I_{j\Delta}^k\right)^2 \\
&= T^{-1}\left[\sum_{k=1}^{K-1}(\tau_{k+1}-\tau_k)\left(\sum_{j=1}^{k}\phi_j\right)^2 + (T-\tau_K)\left(\sum_{j=1}^{K}\phi_j\right)^2\right] \\
&= \begin{cases} (1-r_1)\,\phi_1^2 & \text{if } K=1 \\ \sum_{k=1}^{K-1}(r_{k+1}-r_k)\left(\sum_{j=1}^{k}\phi_j\right)^2 + (1-r_K)\left(\sum_{j=1}^{K}\phi_j\right)^2 & \text{if } K>1. \end{cases}
\end{aligned}$$

The second term, from the proof Lemma 2.1 in the online supplement,

$$T^{-1}\sum_{j=1}^{T}\left[\sigma\sqrt{\Delta}\sum_{i=1}^{j}\varepsilon_{i\Delta}+y_0\right]^2 \Longrightarrow \sigma^2 N\Psi_3.$$

The third term

$$\begin{aligned}
& 2T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\left(\sigma\sqrt{\Delta}\sum_{j=1}^{i}\varepsilon_{j\Delta}+y_0\right) \\
= \;& 2\sigma N^{1/2}\left(T^{-3/2}\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}\sum_{j=1}^{i} I_{j\Delta}^k\sum_{j=1}^{i}\varepsilon_{j\Delta}\right) + 2y_0\left(T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right) \\
\Longrightarrow \;& 2\sigma N^{1/2}\sum_{k=1}^{K}\phi_k\int_{r_k}^{1} w_s ds + 2y_0\sum_{k=1}^{K}\phi_k(1-r_k)
\end{aligned}$$

since

$$\begin{aligned}
& T^{-3/2}\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k\sum_{j=1}^{i} I_{j\Delta}^k\right)\left(\sum_{j=1}^{i}\varepsilon_{j\Delta}\right) \\
= \;& T^{-3/2}\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}\left(\sum_{j=1}^{i} I_{j\Delta}^k\sum_{j=1}^{i}\varepsilon_{j\Delta}\right) \\
= \;& T_n^{-3/2}\sum_{k=1}^{K}\phi_k\left[(T-\tau_k)\sum_{i=1}^{\tau_k}\varepsilon_{i\Delta} + \sum_{i=\tau_k+1}^{T}(T-i+1)\,\varepsilon_{i\Delta}\right] \\
= \;& \sum_{k=1}^{K}\phi_k\left[\frac{T-\tau_k}{T}\left(T^{-1/2}\sum_{i=1}^{\tau_k}\varepsilon_{i\Delta}\right) + T_n^{-3/2}\sum_{i=\tau_k+1}^{T_n}(T-i)\,\varepsilon_{i\Delta}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{K}\phi_k\left[\frac{T-\tau_k}{T}\left(T^{-1/2}\sum_{i=1}^{\tau_k}\varepsilon_{i\Delta}\right)+T^{-1/2}\sum_{i=\tau_k+1}^{T}\varepsilon_{i\Delta}-T^{-3/2}\sum_{i=\tau_k+1}^{T}i\varepsilon_{i\Delta}\right] \\
&\Rightarrow \sum_{k=1}^{K}\phi_k\left[(1-r_k)\,w_{r_k}+(w_1-w_{r_k})-\left(w_1-r_k w_{r_k}-\int_{r_k}^{1}w_s ds\right)\right] \\
&= \sum_{k=1}^{K}\phi_k\int_{r_k}^{1}w_s ds.
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
T^{-1}\sum_{i=1}^{T}y_{i\Delta}^2 \;\Rightarrow\; & \sigma^2 N\left[\Psi_3+\Delta_1+(1-r_K)\left(\sum_{j=1}^{K}\frac{\phi_j}{\sigma N^{1/2}}\right)^2\right. \\
& \left.+2\sum_{k=1}^{K}\frac{\phi_k}{\sigma N^{1/2}}\int_{r_k}^{1}w_s ds+2\frac{y_0}{\sigma N^{1/2}}\sum_{k=1}^{K}\frac{\phi_k}{\sigma N^{1/2}}(1-r_k)\right]\equiv\sigma^2 N\tilde{\Psi}_3
\end{aligned}
$$

with

$$
\Delta_1=\begin{cases}(1-r_1)\frac{\phi_1^2}{\sigma^2 N} & \text{if } K=1\\ \sum_{k=1}^{K-1}(r_{k+1}-r_k)\left(\sum_{j=1}^{k}\frac{\phi_j}{\sigma N^{1/2}}\right)^2+(1-r_K)\left(\sum_{j=1}^{K}\frac{\phi_j}{\sigma N^{1/2}}\right)^2 & \text{if } K>1.\end{cases}
$$

(d) By squaring (9), subtracting $y_{(i-1)\Delta}^2$ from both sides, summing over $i=1,...,T$, re-organizing the equation, and multiplying by $T^{-1/2}$, we get

$$
\begin{aligned}
T^{-1/2}\sum_{i=1}^{T}y_{(i-1)\Delta}\varepsilon_{i\Delta} \;=\; & \frac{T^{-1/2}}{2\sigma\sqrt{\Delta}}\left[y_{T\Delta}^2-y_0^2-\sigma^2\Delta\sum_{i=1}^{T}\varepsilon_{i\Delta}^2-\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\right)^2\right. \\
& \left.-2\sum_{i=1}^{T}\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\left(y_{(i-1)\Delta}+\sigma\sqrt{\Delta}\varepsilon_{(i-1)\Delta}\right)\right].
\end{aligned}
$$

We have $y_{T\Delta}\Longrightarrow\tilde{\Psi}_1, \sigma^2\Delta\sum_{i=1}^{T}\varepsilon_{i\Delta}^2\to\sigma^2 N$, $\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\right)^2=\sum_{k=1}^{K}\phi_k^2$,

$$
\begin{aligned}
& 2\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\right)\left(y_{(i-1)\Delta}+\sigma\sqrt{\Delta}\varepsilon_{(i-1)\Delta}\right) \\
=\; & 2\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}I_{i\Delta}^k y_{(i-1)\Delta}+\sigma\sqrt{\Delta}\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}I_{i\Delta}^k\varepsilon_{(i-1)\Delta} \\
=\; & 2\sum_{k=1}^{K}\phi_k y_{(\tau_k-1)\Delta}+\sigma\sqrt{\Delta}\sum_{k=1}^{K}\phi_k\varepsilon_{(\tau_k-1)\Delta}\Longrightarrow 2\sigma N^{1/2}\sum_{k=1}^{K}\phi_k\left(w_{r_k}+\frac{y_0}{\sigma N^{1/2}}\right)
\end{aligned}
$$

since $\sigma\sqrt{\Delta}\sum_{k=1}^{K}\phi_k\varepsilon_{(\tau_k-1)\Delta}\to 0$ and

$$
y_{(\tau_{k-1})\Delta}=\sigma\sqrt{\Delta}\sum_{j=1}^{\tau_{k-1}}\varepsilon_j+y_0+\sum_{k=1}^{K}\phi_k\sum_{j=1}^{\tau_{k-1}}I_j^k\Longrightarrow N^{1/2}\left(w_{r_k}+\frac{y_0}{\sigma N^{1/2}}+\Delta_2\right),
$$

where $\Delta_2 = 0$ if $K = 0$ and $\Delta_2 = \sum_{j=1}^{k-1} \frac{\phi_j}{\sigma N^{1/2}}$ if $K > 1$. Therefore,

$$\begin{aligned} T^{-1/2}\sum_{i=1}^{T} y_{(i-1)\Delta}\varepsilon_{i\Delta} &\Longrightarrow \frac{\sigma N^{1/2}}{2}\left[\tilde{\Psi}_1^2 - \frac{y_0^2}{\sigma^2 N} - 1 - \sum_{k=1}^{K}\frac{\phi_k^2}{\sigma^2 N} - 2\sum_{k=1}^{K}\frac{\phi_k}{\sigma N^{1/2}}\left(w_{r_k} + \frac{y_0}{\sigma N^{1/2}} + \Delta_2\right)\right] \\ &\equiv \sigma N^{1/2}\tilde{\Psi}_4. \end{aligned}$$

■

Proof of Lemma 3.2. (a) The alternative model can be rewritten as

$$y_{i\Delta} = \alpha_0 \frac{1 - e^{i\theta\Delta}}{1 - e^{\theta\Delta}} + \sum_{j=1}^{i} e^{(i-j)\theta\Delta}\left(\lambda_0\varepsilon_{j\Delta} + \sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right) + e^{i\theta\Delta}y_0. \tag{26}$$

It follows that

$$y_{T\Delta} = \alpha_0 \frac{1 - e^{T\theta\Delta}}{1 - e^{\theta\Delta}} + \lambda_0\sum_{j=1}^{T} e^{(T-j)\theta\Delta}\varepsilon_{j\Delta} + e^{T\theta\Delta}y_0 + \sum_{j=1}^{T} e^{(T-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k.$$

From the proof of Lemma 2.2 in the online supplement

$$\alpha_0 \frac{1 - e^{T\theta\Delta}}{1 - e^{\theta\Delta}} + \lambda_0\sum_{j=1}^{T} e^{(T-j)\theta\Delta}\varepsilon_{j\Delta} + e^{T\theta\Delta}y_0 \Longrightarrow \sigma N^{1/2}\Xi_1.$$

The last term

$$\sum_{j=1}^{T} e^{(T-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k = \sum_{k=1}^{K}\phi_k\sum_{j=1}^{T} e^{\frac{T-j}{T}c} I_{j\Delta}^k = \sum_{k=1}^{K}\phi_k e^{(1-r_k)c}.$$

Therefore,

$$y_{T\Delta} \Longrightarrow \sigma N^{1/2}\left[\Xi_1 + \sum_{k=1}^{K}\frac{\phi_k}{\sigma N^{1/2}}e^{(1-r_k)c}\right] \equiv \sigma N^{1/2}\tilde{\Xi}_1.$$

(b) The quantity

$$\begin{aligned} T^{-1}\sum_{i=1}^{T} y_{i\Delta} &= T^{-1}\sum_{i=1}^{T}\left[\alpha_0\frac{1 - e^{i\theta\Delta}}{1 - e^{\theta\Delta}} + \lambda_0\sum_{j=1}^{i} e^{(T-j)\theta\Delta}\varepsilon_{j\Delta} + e^{i\theta\Delta}y_0 + \sum_{j=1}^{i} e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right] \\ &= T^{-1}\sum_{i=1}^{T}\left[\alpha_0\frac{1 - e^{i\theta\Delta}}{1 - e^{\theta\Delta}} + \lambda_0\sum_{j=1}^{i} e^{(T-j)\theta\Delta}\varepsilon_{j\Delta} + e^{i\theta\Delta}y_0\right] + T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i} e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k. \end{aligned}$$

From the proof of Lemma 2.2 in the online supplement

$$T^{-1}\sum_{i=1}^{T}\left[\alpha_0\frac{1 - e^{i\theta\Delta}}{1 - e^{\theta\Delta}} + \lambda_0\sum_{j=1}^{i} e^{(T-j)\theta\Delta}\varepsilon_{j\Delta} + e^{i\theta\Delta}y_0\right] \Longrightarrow \sigma N^{1/2}\Xi_2.$$

Furthermore,

$$T^{-1}\sum_{i=1}^{T}\sum_{j=1}^{i} e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k \quad = \quad T^{-1}\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}\sum_{j=1}^{i} e^{(i-j)\theta\Delta} I_{j\Delta}^k$$

$$
\begin{aligned}
&= T^{-1}\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}\frac{1-e^{(T-i+1)\theta\Delta}}{1-e^{\theta\Delta}}I_{i\Delta}^{k} \\
&= T^{-1}\sum_{k=1}^{K}\phi_k\frac{1-e^{(T-\tau_k+1)\theta\Delta}}{1-e^{\theta\Delta}} \\
&\to \frac{1}{c}\sum_{k=1}^{K}\phi_k\left[e^{(1-r_k)c}-1\right].
\end{aligned}
$$

Therefore,

$$
T^{-1}\sum_{i=1}^{T}y_{i\Delta}\Longrightarrow \sigma N^{1/2}\left\{\Xi_2+\frac{1}{c}\sum_{k=1}^{K}\frac{\phi_k}{\sigma N^{1/2}}\left[e^{(1-r_k)c}-1\right]\right\}\equiv \sigma N^{1/2}\tilde{\Xi}_2.
$$

(c) The quantity

$$
\begin{aligned}
T^{-1}\sum_{i=1}^{T}y_{i\Delta}^2 =& T^{-1}\sum_{i=1}^{T}\left[\alpha_0\frac{1-e^{i\theta\Delta}}{1-e^{\theta\Delta}}+\lambda_0\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\varepsilon_{j\Delta}+e^{i\theta\Delta}y_0+\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k}\right]^2 \\
&= T^{-1}\sum_{i=1}^{T}\left(\alpha_0\frac{1-e^{i\theta\Delta}}{1-e^{\theta\Delta}}+\lambda_0\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\varepsilon_{j\Delta}+e^{i\theta\Delta}y_0\right)^2+T^{-1}\sum_{i=1}^{T}\left(\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k}\right)^2 \\
&+2T^{-1}\sum_{i=1}^{T}\left(\alpha_0\frac{1-e^{i\theta\Delta}}{1-e^{\theta\Delta}}\right)\left(\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k}\right) \\
&+2T^{-1}\sum_{i=1}^{T}\lambda_0\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\varepsilon_{j\Delta}\left(\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k}\right) \\
&+2T^{-1}\sum_{i=1}^{T}e^{i\theta\Delta}y_0\left(\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k}\right).
\end{aligned}
$$

From the proof of Lemma 2.2 in the online supplement

$$
T^{-1}\sum_{i=1}^{T}\left(\alpha_0\frac{1-e^{i\theta\Delta}}{1-e^{\theta\Delta}}+\lambda_0\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\varepsilon_{j\Delta}+e^{i\theta\Delta}y_0\right)^2\Longrightarrow \sigma^2 N\Xi_3.
$$

The second term

$$
T^{-1}\sum_{i=1}^{T}\left(\sum_{j=1}^{i}e^{(i-j)\theta\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^{k}\right)^2=T^{-1}\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k\sum_{j=1}^{i}e^{(i-j)\theta\Delta}I_{j\Delta}^{k}\right)^2.
$$

If $K=1$, we have

$$
T^{-1}\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k\sum_{j=1}^{i}e^{(i-j)\theta\Delta}I_{j\Delta}^{k}\right)^2=T^{-1}\sum_{i=\tau_1}^{T}\phi_1^2e^{2(i-\tau_1)\theta\Delta}=\frac{\Delta}{N}\phi_1^2\frac{e^{2c(1-r_1)}-1}{e^{2\theta\Delta}-1}\to\phi_1^2\frac{1}{2c}\left[e^{2c(1-r_1)}-1\right].
$$

If $K > 1$,

$$\begin{aligned}
& T^{-1} \sum_{i=1}^{T} \left(\sum_{k=1}^{K} \phi_k \sum_{j=1}^{i} e^{(i-j)\theta\Delta} I_{j\Delta}^k \right)^2 \\
= \;& T^{-1} \sum_{k=1}^{K-1} \sum_{i=\tau_k}^{\tau_{k+1}-1} \left(\sum_{j=1}^{k} \phi_j e^{(i-\tau_j)\theta\Delta} \right)^2 + T^{-1} \sum_{i=\tau_K}^{T} \left(\sum_{j=1}^{K} \phi_j e^{(i-\tau_j)\theta\Delta} \right)^2 \\
= \;& T^{-1} \sum_{k=1}^{K-1} \sum_{i=\tau_k}^{\tau_{k+1}-1} e^{2i\theta\Delta} \left(\sum_{j=1}^{k} \phi_j e^{-\tau_j\theta\Delta} \right)^2 + T^{-1} \sum_{i=\tau_K}^{T} e^{2i\theta\Delta} \left(\sum_{j=1}^{K} \phi_j e^{-\tau_j\theta\Delta} \right)^2 \\
= \;& T^{-1} \sum_{k=1}^{K-1} \left(\sum_{j=1}^{k} \phi_j e^{-\tau_j\theta\Delta} \right)^2 \left(\sum_{i=\tau_k}^{\tau_{k+1}-1} e^{2i\theta\Delta} \right) + T^{-1} \left(\sum_{j=1}^{K} \phi_j e^{-\tau_j\theta\Delta} \right)^2 \sum_{i=\tau_K}^{T} e^{2i\theta\Delta} \\
= \;& T^{-1} \sum_{k=1}^{K-1} \left(\sum_{j=1}^{k} \phi_j e^{-\tau_j\theta\Delta} \right)^2 \left[e^{2r_k\theta} \frac{1 - e^{2\theta(r_{k+1}-r_k)}}{1 - e^{2\theta\Delta}} \right] + T^{-1} \left(\sum_{j=1}^{K} \phi_j e^{-\tau_j\theta\Delta} \right)^2 \left[e^{2r_K\theta} \frac{1 - e^{2\theta(1-r_K)}}{1 - e^{2\theta\Delta}} \right] \\
\to \;& \frac{1}{2c} \left[\sum_{k=1}^{K-1} \left(\sum_{j=1}^{k} \phi_j e^{-r_j\theta} \right)^2 e^{2r_k\theta} \left[e^{2\theta(r_{k+1}-r_k)} - 1 \right] + \left(\sum_{j=1}^{K} \phi_j e^{-r_j\theta} \right)^2 e^{2r_K\theta} \left[e^{2\theta(1-r_K)} - 1 \right] \right].
\end{aligned}$$

The third term

$$\begin{aligned}
& 2T^{-1} \sum_{i=1}^{T} \left(\alpha_0 \frac{1 - e^{i\theta\Delta}}{1 - e^{\theta\Delta}} \right) \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} \sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right) \\
= \;& 2 \frac{\alpha_0}{1 - e^{\theta\Delta}} T^{-1} \sum_{i=1}^{T} \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} \sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right) - 2 \frac{\alpha_0}{1 - e^{\theta\Delta}} T^{-1} \sum_{i=1}^{T} e^{i\theta\Delta} \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} \sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right) \\
= \;& 2\mu T^{-1} \sum_{k=1}^{K} \phi_k \sum_{i=1}^{T} \sum_{j=1}^{i} e^{(i-j)\theta\Delta} I_{j\Delta}^k - 2\mu T^{-1} \sum_{k=1}^{K} \phi_k \sum_{i=1}^{T} \sum_{j=1}^{i} e^{i\theta\Delta} e^{(i-j)\theta\Delta} I_{j\Delta}^k \\
= \;& 2\mu T^{-1} \sum_{k=1}^{K} \phi_k e^{-\tau_k\theta\Delta} \sum_{i=\tau_k}^{T} e^{i\theta\Delta} - 2\mu T^{-1} \sum_{k=1}^{K} \phi_k e^{-\tau_k\theta\Delta} \sum_{i=\tau_k}^{T} e^{2i\theta\Delta} \\
= \;& 2\mu T^{-1} \sum_{k=1}^{K} \phi_k e^{-\tau_k\theta\Delta} \left[\sum_{i=\tau_k}^{T} e^{i\theta\Delta} - \sum_{i=\tau_k}^{T} e^{2i\theta\Delta} \right] \\
= \;& \frac{\mu}{c} \sum_{k=1}^{K} \phi_k \left[2e^{c(1-r_k)} - 2 - e^{c(2-r_k)} + e^{r_k c} \right].
\end{aligned}$$

The fourth term

$$\begin{aligned}
& 2T^{-1} \sum_{i=1}^{T} \lambda_0 \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} \varepsilon_{j\Delta} \right) \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} \sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right) \\
= \;& 2\lambda_0 T^{-1} \sum_{k=1}^{K} \phi_k \sum_{i=1}^{T} \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} \varepsilon_{j\Delta} \right) \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} I_{j\Delta}^k \right)
\end{aligned}$$

$$= \quad 2\lambda_0 T^{1/2} \sum_{k=1}^{K} \phi_k \left[\frac{1}{T} \sum_{i=\tau_k}^{T} e^{(i-\tau_k)\theta\Delta} \left(T^{-1/2} \sum_{j=1}^{i} e^{(i-j)\theta\Delta} \varepsilon_{j\Delta} \right) \right]$$

$$\Longrightarrow \quad 2\sigma N^{1/2} \sum_{k=1}^{K} \phi_k \int_{r_k}^{1} e^{c(r-r_k)} J_c(r)\, dr.$$

The fifth term

$$\begin{aligned} 2T^{-1} \sum_{i=1}^{T} e^{i\theta\Delta} y_0 \left(\sum_{j=1}^{i} e^{(i-j)\theta\Delta} \sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right) &= 2y_0 T^{-1} \sum_{k=1}^{K} \phi_k \sum_{i=1}^{T} \left(\sum_{j=1}^{i} e^{i\theta\Delta} e^{(i-j)\theta\Delta} I_{j\Delta}^k \right) \\ &= 2y_0 T^{-1} \sum_{k=1}^{K} \phi_k \sum_{i=\tau_k}^{T} e^{(2i-\tau_k)\theta\Delta} \\ &= 2y_0 \sum_{k=1}^{K} \phi_k \left(e^{-\tau_k\theta\Delta} T^{-1} \sum_{i=\tau_k}^{T} e^{2i\theta\Delta} \right) \\ &= \frac{y_0}{c} \sum_{k=1}^{K} \phi_k e^{r_k c} \left[e^{2c(1-r_k)} - 1 \right]. \end{aligned}$$

Therefore,

$$\begin{aligned} T^{-1} \sum_{i=1}^{T} y_{(i-1)\Delta}^2 &\Longrightarrow \sigma^2 N \Xi_3 + \frac{1}{2c} \sum_{k=1}^{K} \phi_k^2 \left[e^{2(1-r_k)c} - 1 \right] + \frac{\mu}{c} \sum_{k=1}^{K} \phi_k \left[2e^{c(1-r_k)} - 2 - e^{c(2-r_k)} + e^{r_k c} \right] \\ &+ 2\sigma N^{1/2} \sum_{k=1}^{K} \phi_k \int_{r_k}^{1} e^{c(r-r_k)} J_c(r)\, dr + \frac{y_0}{c} \sum_{k=1}^{K} \phi_k e^{r_k c} \left[e^{2c(1-r_k)} - 1 \right] \\ &= \sigma^2 N \left\{ \Xi_3 + \Delta_3 + \frac{1}{c} \frac{\mu}{\sigma N^{1/2}} \sum_{k=1}^{K} \frac{\phi_k}{\sigma N^{1/2}} \left[2e^{c(1-r_k)} - 2 - e^{c(2-r_k)} + e^{r_k c} \right] \right. \\ &\left. + 2 \sum_{k=1}^{K} \frac{\phi_k}{\sigma N^{1/2}} \int_{r_k}^{1} e^{c(r-r_k)} J_c(r)\, dr + \frac{1}{c} \frac{y_0}{\sigma N^{1/2}} \sum_{k=1}^{K} \frac{\phi_k}{\sigma N^{1/2}} e^{r_k c} \left[e^{2c(1-r_k)} - 1 \right] \right\} \\ &\equiv \sigma^2 N \tilde{\Xi}_3, \end{aligned}$$

where

$$\Delta_3 = \begin{cases} \frac{\phi_1^2}{\sigma^2 N} \frac{1}{2c} \left[e^{2c(1-r_1)} - 1 \right] & \text{if } K = 1 \\ \frac{1}{2c} \left[\sum_{k=1}^{K-1} \left(\sum_{j=1}^{k} \frac{\phi_j}{\sigma N^{1/2}} e^{-r_j\theta} \right)^2 e^{2r_k\theta} \left[e^{2\theta(r_{k+1}-r_k)} - 1 \right] + \left(\sum_{j=1}^{K} \frac{\phi_j}{\sigma N^{1/2}} e^{-r_j\theta} \right)^2 e^{2r_K\theta} \left[e^{2\theta(1-r_K)} - 1 \right] \right] & \text{if } K > 1 \end{cases}.$$

(d) By squaring (10), subtracting $y_{(i-1)\Delta}^2$ from both sides, summing over $i = 1, ..., T$, re-organizing the equation, and multiplying T^{-1} such that

$$T^{-1/2} \sum_{i=1}^{T} y_{(i-1)\Delta} \varepsilon_{i\Delta}$$

$$= \frac{T^{-1/2}}{2e^{\theta\Delta}\lambda_0} \left[\sum_{i=1}^{T} \left(y_{i\Delta}^2 - y_{(i-1)\Delta}^2 \right) - T\alpha_0^2 - \left(e^{2\theta\Delta} - 1 \right) \sum_{i=1}^{T} y_{(i-1)\Delta}^2 - \lambda_0^2 \sum_{i=1}^{T} \varepsilon_{i\Delta}^2 - 2\alpha_0 e^{\theta\Delta} \sum_{i=1}^{T} y_{(i-1)\Delta} \right.$$

$$-2\alpha_0\lambda_0\sum_{i=1}^{T}\varepsilon_{i\Delta}-\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\right)^2-2\alpha_0\sum_{i=1}^{T}\sum_{k=1}^{K}\phi_k I_{i\Delta}^k-2e^{\theta\Delta}\sum_{i=1}^{T}y_{(i-1)\Delta}\sum_{k=1}^{K}\phi_k I_{i\Delta}^k-2\lambda_0\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\right)\varepsilon_{i\Delta}\Bigg].$$

From (a), (b) and (c),

$$\frac{T^{-1/2}}{2e^{\theta\Delta}\lambda_0}\left[\sum_{i=1}^{T}\left(y_{i\Delta}^2-y_{(i-1)\Delta}^2\right)-T\alpha_0^2-\left(e^{2\theta\Delta}-1\right)\sum_{i=1}^{T}y_{(i-1)\Delta}^2-\lambda_0^2\sum_{i=1}^{T}\varepsilon_{i\Delta}^2-2\alpha_0e^{\theta\Delta}\sum_{i=1}^{T}y_{(i-1)\Delta}-2\alpha_0\lambda_0\sum_{i=1}^{T}\varepsilon_{i\Delta}\right]$$
$$\Longrightarrow\frac{\sigma N^{1/2}}{2}\left[\tilde{\Xi}_1^2-\frac{y_0^2}{\sigma^2N}-2c\tilde{\Xi}_3-1+2c\frac{\mu}{\sigma N^{1/2}}\tilde{\Xi}_2\right].$$

Moreover,

$$\begin{aligned}
\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\right)^2 &= \sum_{k=1}^{K}\phi_k^2\sum_{i=1}^{T}I_{i\Delta}^k=\sum_{k=1}^{K}\phi_k^2;\\
2\alpha_0\sum_{i=1}^{T}\sum_{k=1}^{K}\phi_k I_{i\Delta}^k &= 2\mu\left(1-e^{\theta\Delta}\right)\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}I_{i\Delta}^k=2\mu\left(1-e^{\theta\Delta}\right)\sum_{k=1}^{K}\phi_k\to 0;\\
2\lambda_0\sum_{i=1}^{T}\left(\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\right)\varepsilon_{i\Delta} &= 2\lambda_0\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}I_{i\Delta}^k\varepsilon_{i\Delta}=2\lambda_0\sum_{k=1}^{K}\phi_k\varepsilon_{\tau_k\Delta}\to 0;
\end{aligned}$$

and

$$\begin{aligned}
&2e^{\theta\Delta}\sum_{i=1}^{T}y_{(i-1)\Delta}\sum_{k=1}^{K}\phi_k I_{i\Delta}^k\\
=\;&2e^{\theta\Delta}\sum_{k=1}^{K}\phi_k\sum_{i=1}^{T}y_{(i-1)\Delta}I_{i\Delta}^k=2e^{\theta\Delta}\sum_{k=1}^{K}\phi_k y_{(\tau_k-1)\Delta}\\
=\;&2e^{\theta\Delta}\sum_{k=1}^{K}\phi_k\left[\alpha_0\frac{1-e^{(\tau_k-1)\theta\Delta}}{1-e^{\theta\Delta}}+\lambda_0\sum_{j=1}^{\tau_k-1}e^{(\tau_k-1-j)\theta\Delta}\varepsilon_{j\Delta}+e^{(\tau_k-1)\theta\Delta}y_0+\sum_{j=1}^{\tau_k-1}e^{(\tau_k-1-j)\theta\Delta}\sum_{s=1}^{K}\phi_s I_{j\Delta}^s\right]\\
\Longrightarrow\;&2\sigma^2N\sum_{k=1}^{K}\frac{\phi_k}{N^{1/2}\sigma}\left[\frac{\mu}{N^{1/2}\sigma}\left(1-e^{r_kc}\right)+J_c\left(r_k\right)+e^{r_kc}\frac{y_0}{N^{1/2}\sigma}+\Delta_4\right],
\end{aligned}\tag{27}$$

where $\Delta_4=0$ if $K=1$ and $\Delta_4=\sum_{j=1}^{k-1}e^{(r_k-r_j)\theta}\frac{\phi_j}{N^{1/2}\sigma}$ if $K>1$. This is because

$$\alpha_0\frac{1-e^{(\tau_k-1)\theta\Delta}}{1-e^{\theta\Delta}}+\lambda_0\sum_{j=1}^{\tau_k-1}e^{(\tau_k-1-j)\theta\Delta}\varepsilon_{j\Delta}+e^{(\tau_k-1)\theta\Delta}y_0\Longrightarrow N^{1/2}\sigma\left[\frac{\mu}{N^{1/2}\sigma}\left(1-e^{r_kc}\right)+J_c\left(r_k\right)+e^{r_kc}\frac{y_0}{N^{1/2}\sigma}\right]$$

and

$$\sum_{j=1}^{\tau_k-1}e^{(\tau_k-1-j)\theta\Delta}\sum_{s=1}^{K}\phi_s I_{j\Delta}^s=\begin{cases}\sum_{s=1}^{K}\phi_s e^{(\tau_k-1)\theta\Delta}\sum_{j=1}^{\tau_k-1}e^{-j\theta\Delta}I_{j\Delta}^s=0 & \text{if } K=1\\ \sum_{j=1}^{k-1}e^{(\tau_k-1-\tau_j)\theta\Delta}\phi_j\;\to N^{1/2}\sigma\sum_{j=1}^{k-1}e^{(r_k-r_j)\theta}\frac{\phi_j}{N^{1/2}\sigma} & \text{if } K>1\end{cases}.$$

Therefore,

$$T^{-1/2}\sum_{i=1}^{T}y_{(i-1)\Delta}\varepsilon_{i\Delta}=\frac{\sigma N^{1/2}}{2}\Bigg\{\tilde{\Xi}_1^2-\frac{y_0^2}{\sigma^2N}-2c\tilde{\Xi}_3-1+2c\frac{\mu}{\sigma N^{1/2}}\tilde{\Xi}_2-\sum_{k=1}^{K}\frac{\phi_k^2}{\sigma^2N}$$

$$-2\sum_{k=1}^{K}\frac{\phi_k}{N^{1/2}\sigma}\left[\frac{\mu}{N^{1/2}\sigma}\left(1-e^{r_kc}\right)+J_c\left(r_k\right)+e^{r_kc}\frac{y_0}{N^{1/2}\sigma}+\Delta_4\right]\Bigg\}$$
$$=\sigma N^{1/2}\tilde{\Xi}_4.$$

■

Appendix C: the DF^J Test Statistic

The least square estimators of the standardized intercept and the autoregressive coefficients with regression (11) are

$$\begin{bmatrix}\tilde{\alpha}-\alpha_0\\ \tilde{\phi}_1-\phi_1\\ \vdots\\ \tilde{\phi}_K-\phi_K\\ \tilde{\beta}-\beta_0\end{bmatrix}=\sigma\sqrt{\Delta}\begin{bmatrix}T & \sum I_{i\Delta}^1 & \cdots & \sum I_{i\Delta}^K & \sum y_{(i-1)\Delta}\\ \sum I_{i\Delta}^1 & \sum\left(I_{i\Delta}^1\right)^2 & \cdots & \sum I_{i\Delta}^1 I_{i\Delta}^K & \sum I_{i\Delta}^1 y_{(i-1)\Delta}\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ \sum I_{i\Delta}^K & \sum I_{i\Delta}^K I_{i\Delta}^1 & \cdots & \sum\left(I_{i\Delta}^K\right)^2 & \sum I_{i\Delta}^K y_{(i-1)\Delta}\\ \sum y_{(i-1)\Delta} & \sum y_{(i-1)\Delta} I_{i\Delta}^1 & \cdots & \sum y_{(i-1)\Delta} I_{i\Delta}^K & \sum y_{(i-1)\Delta}^2\end{bmatrix}^{-1}\begin{bmatrix}\sum\varepsilon_{i\Delta}\\ \sum I_{i\Delta}^1\varepsilon_{i\Delta}\\ \vdots\\ \sum I_{i\Delta}^K\varepsilon_{i\Delta}\\ \sum y_{(i-1)\Delta}\varepsilon_{i\Delta}\end{bmatrix}.$$

Let $\sum$ denote summation over $i=1,\cdots,T$. Based on Lemma 3.1(d) and 3.2(d), the appropriate scaling matrix is $\Upsilon_{T_n}=diag\left(\sqrt{T},1,\cdots,1,\sqrt{T}\right)$. Pre-multiplying the above equation by Υ_T leads to

$$T^{1/2}\Upsilon_T\begin{bmatrix}\tilde{\alpha}-\alpha_0\\ \tilde{\phi}_1-\phi_1\\ \vdots\\ \tilde{\phi}_K-\phi_K\\ \tilde{\beta}-\beta_0\end{bmatrix}=\sigma N^{1/2}\left\{\Upsilon_T^{-1}\begin{bmatrix}T & \sum I_{i\Delta}^1 & \cdots & \sum I_{i\Delta}^K & \sum y_{(i-1)\Delta}\\ \sum I_{i\Delta}^1 & \sum\left(I_{i\Delta}^1\right)^2 & \cdots & \sum I_{i\Delta}^1 I_{i\Delta}^K & \sum I_{i\Delta}^1 y_{(i-1)\Delta}\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ \sum I_{i\Delta}^K & \sum I_{i\Delta}^K I_{i\Delta}^1 & \cdots & \sum\left(I_{i\Delta}^K\right)^2 & \sum I_{i\Delta}^K y_{(i-1)\Delta}\\ \sum y_{(i-1)\Delta} & \sum y_{(i-1)\Delta} I_{i\Delta}^1 & \cdots & \sum y_{(i-1)\Delta} I_{i\Delta}^K & \sum y_{(i-1)\Delta}^2\end{bmatrix}\Upsilon_T^{-1}\right\}^{-1}$$
$$\Upsilon_T^{-1}\begin{bmatrix}\sum\varepsilon_{i\Delta}\\ \sum I_{i\Delta}^1\varepsilon_{i\Delta}\\ \vdots\\ \sum I_{i\Delta}^K\varepsilon_{i\Delta}\\ \sum y_{(i-1)\Delta}\varepsilon_{i\Delta}\end{bmatrix}$$

with $\alpha_0=0$ and $\beta_0=1$ under the null. The first term

$$\Upsilon_T^{-1}\begin{bmatrix}T & \sum I_{i\Delta}^1 & \cdots & \sum I_{i\Delta}^K & \sum y_{(i-1)\Delta}\\ \sum I_{i\Delta}^1 & \sum\left(I_{i\Delta}^1\right)^2 & \cdots & \sum I_{i\Delta}^1 I_{i\Delta}^K & \sum I_{i\Delta}^1 y_{(i-1)\Delta}\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ \sum I_{i\Delta}^K & \sum I_{i\Delta}^K I_{i\Delta}^1 & \cdots & \sum\left(I_{i\Delta}^K\right)^2 & \sum I_{i\Delta}^K y_{(i-1)\Delta}\\ \sum y_{(i-1)\Delta} & \sum y_{(i-1)\Delta} I_{i\Delta}^1 & \cdots & \sum y_{(i-1)\Delta} I_{i\Delta}^K & \sum y_{(i-1)\Delta}^2\end{bmatrix}\Upsilon_T^{-1}$$
$$=\begin{bmatrix}1 & T^{-1/2}\sum I_{i\Delta}^1 & \cdots & T^{-1/2}\sum I_{i\Delta}^K & T^{-1}\sum y_{(i-1)\Delta}\\ T^{-1/2}\sum I_{i\Delta}^1 & \sum\left(I_{i\Delta}^1\right)^2 & \cdots & \sum I_{i\Delta}^1 I_{i\Delta}^K & T^{-1/2}\sum I_{i\Delta}^1 y_{(i-1)\Delta}\\ \vdots & \vdots & \cdots & \vdots & \vdots\\ T^{-1/2}\sum I_{i\Delta}^K & \sum I_{i\Delta}^K I_{i\Delta}^1 & \cdots & \sum\left(I_{i\Delta}^K\right)^2 & T^{-1/2}\sum I_{i\Delta}^K y_{(i-1)\Delta}\\ T^{-1}\sum y_{(i-1)\Delta} & T^{-1/2}\sum y_{(i-1)\Delta} I_{i\Delta}^1 & \cdots & T^{-1/2}\sum y_{(i-1)\Delta} I_{i\Delta}^K & T^{-1}\sum y_{(i-1)\Delta}^2\end{bmatrix}$$

$$= \begin{bmatrix} 1 & T^{-1/2} & \cdots & T^{-1/2} & T^{-1}\sum y_{(i-1)\Delta} \\ T^{-1/2} & 1 & \cdots & 0 & T^{-1/2}y_{(\tau_1-1)\Delta} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ T^{-1/2} & 0 & \cdots & 1 & T^{-1/2}y_{(\tau_K-1)\Delta} \\ T^{-1}\sum y_{(i-1)\Delta} & T^{-1/2}y_{(\tau_1-1)\Delta} & \cdots & T^{-1/2}y_{(\tau_K-1)\Delta} & T^{-1}\sum y_{(i-1)\Delta}^2 \end{bmatrix}.$$

The second term

$$\Upsilon_T^{-1}\begin{bmatrix} \sum \varepsilon_{i\Delta} \\ \sum I_{i\Delta}^1 \varepsilon_{i\Delta} \\ \vdots \\ \sum I_{i\Delta}^K \varepsilon_{i\Delta} \\ \sum y_{(i-1)\Delta}\varepsilon_{i\Delta} \end{bmatrix} = \begin{bmatrix} T^{-1/2}\sum \varepsilon_{i\Delta} \\ \sum I_{i\Delta}^1 \varepsilon_{i\Delta} \\ \vdots \\ \sum I_{i\Delta}^K \varepsilon_{i\Delta} \\ T^{-1/2}\sum y_{(i-1)\Delta}\varepsilon_{i\Delta} \end{bmatrix}.$$

Proof of Theorem 3.1. Under the null hypothesis of (9), from Lemma 3.1, the first term

$$\Upsilon_T^{-1}\begin{bmatrix} T & \sum I_{i\Delta}^1 & \cdots & \sum I_{i\Delta}^K & \sum y_{(i-1)\Delta} \\ \sum I_{i\Delta}^1 & \sum \left(I_{i\Delta}^1\right)^2 & \cdots & \sum I_{i\Delta}^1 I_{i\Delta}^K & \sum I_{i\Delta}^1 y_{(i-1)\Delta} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \sum I_{i\Delta}^K & \sum I_{i\Delta}^K I_{i\Delta}^1 & \cdots & \sum \left(I_{i\Delta}^K\right)^2 & \sum I_{i\Delta}^K y_{(i-1)\Delta} \\ \sum y_{(i-1)\Delta} & \sum y_{(i-1)\Delta} I_{i\Delta}^1 & \cdots & \sum y_{(i-1)\Delta} I_{i\Delta}^K & \sum y_{(i-1)\Delta}^2 \end{bmatrix}\Upsilon_T^{-1} \Rightarrow \begin{bmatrix} 1 & 0 & \cdots & 0 & \sigma N^{1/2}\tilde{\Psi}_2 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \sigma N^{1/2}\tilde{\Psi}_2 & 0 & \cdots & 0 & \sigma^2 N\tilde{\Psi}_3 \end{bmatrix}$$

and the second term

$$\Upsilon_T^{-1}\begin{bmatrix} \sum \varepsilon_{i\Delta} \\ \sum I_{i\Delta}^1 \varepsilon_{i\Delta} \\ \vdots \\ \sum I_{i\Delta}^K \varepsilon_{i\Delta} \\ \sum y_{(i-1)\Delta}\varepsilon_{i\Delta} \end{bmatrix} \Rightarrow \begin{bmatrix} w_1 \\ \varepsilon_{\tau_1} \\ \vdots \\ \varepsilon_{\tau_K} \\ \sigma N^{1/2}\tilde{\Psi}_4 \end{bmatrix}.$$

Combing these two terms, we have

$$\begin{bmatrix} T\tilde{\alpha} \\ T^{1/2}\left(\tilde{\phi}_1 - \phi_1\right) \\ \vdots \\ T^{1/2}\left(\tilde{\phi}_K - \phi_K\right) \\ T\left(\tilde{\beta}-1\right) \end{bmatrix} \Rightarrow \sigma N^{1/2}\begin{bmatrix} 1 & 0 & \cdots & 0 & \sigma N^{1/2}\tilde{\Psi}_2 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \sigma N^{1/2}\tilde{\Psi}_2 & 0 & \cdots & 0 & \sigma^2 N\tilde{\Psi}_3 \end{bmatrix}^{-1}\begin{bmatrix} w_1 \\ \varepsilon_{\tau_1} \\ \vdots \\ \varepsilon_{\tau_K} \\ \sigma N^{1/2}\tilde{\Psi}_4 \end{bmatrix} = \begin{bmatrix} \sigma N^{1/2}\frac{\tilde{\Psi}_3 w_1 - \tilde{\Psi}_2\tilde{\Psi}_4}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2} \\ \sigma N^{1/2}\varepsilon_{\tau_1} \\ \vdots \\ \sigma N^{1/2}\varepsilon_{\tau_K} \\ \frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2} \end{bmatrix}.$$

Therefore,

$$\begin{aligned} T\tilde{\alpha} &\Rightarrow \sigma N^{1/2}\frac{\tilde{\Psi}_3 w_1 - \tilde{\Psi}_2\tilde{\Psi}_4}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2} \\ T^{1/2}\left(\tilde{\phi}_k - \phi_k\right) &\Rightarrow N\left(0, \sigma^2 N\right) \quad \text{for } k = 1, \ldots, K \\ T\left(\tilde{\beta}-1\right) &\Rightarrow \frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2}. \end{aligned}$$

The estimated error variance

$$\tilde{\sigma}_v^2 = \sum \left(y_{i\Delta} - \tilde{\beta}y_{(i-1)\Delta} - \tilde{\alpha} - \sum_{k=1}^K \tilde{\phi}_k I_{i\Delta}^k\right)^2$$

$$
\begin{aligned}
&= \sum \left[\sigma\sqrt{\Delta}\varepsilon_{i\Delta} - \left(\tilde{\beta}-1\right) y_{(i-1)\Delta} - \tilde{\alpha} + \sum_{k=1}^{K}\left(\tilde{\phi}_k - \phi_k\right) I_{i\Delta}^k\right]^2 \\
&= \sigma^2\Delta \sum \varepsilon_{i\Delta}^2 + \left(\tilde{\beta}-1\right)^2 \sum y_{(i-1)\Delta}^2 + \tilde{\alpha}^2 + \sum\left(\sum_{k=1}^{K}\left(\tilde{\phi}_k - \phi_k\right) I_{i\Delta}^k\right)^2 \\
&\quad -2\sigma\sqrt{\Delta}\left(\tilde{\beta}-1\right)\sum y_{(i-1)\Delta}\varepsilon_{i\Delta} - 2\sigma\sqrt{\Delta}\tilde{\alpha}\sum \varepsilon_{i\Delta} + 2\sigma\sqrt{\Delta}\sum\left(\sum_{k=1}^{K}\left(\tilde{\phi}_k - \phi_k\right) I_{i\Delta}^k\right)\varepsilon_{i\Delta} \\
&\quad +2\left(\tilde{\beta}-1\right)\tilde{\alpha}\sum y_{(i-1)\Delta} - 2\left(\tilde{\beta}-1\right)\sum y_{(i-1)\Delta}\sum_{k=1}^{K}\left(\tilde{\phi}_k - \phi_k\right) I_{i\Delta}^k - 2\tilde{\alpha}\sum\sum_{k=1}^{K}\left(\tilde{\phi}_k - \phi_k\right) I_{i\Delta}^k \\
&= \sigma^2\Delta\sum\varepsilon_{i\Delta}^2\left[1+o_p(1)\right] \to \sigma^2 N
\end{aligned}
$$

from Lemma 3.1 and the fact that $\tilde{\alpha} = O_p\left(T^{-1}\right)$, $\tilde{\beta}-1 = O_p\left(T^{-1}\right)$, and $\tilde{\phi}_k - \phi_k = O_p\left(T^{-1/2}\right)$,

$$
DF^J = \frac{\left(\tilde{\beta}-1\right)\left[T\sum_{j=1}^{T} y_{j\Delta}^2 - \left(\sum_{j=1}^{T} y_{(j-1)\Delta}\right)^2\right]^{1/2}}{\left[\sum_{j=1}^{T}\left(y_{i\Delta} - \tilde{\beta}y_{(i-1)\Delta} - \tilde{\alpha} - \sum_{k=1}^{K}\tilde{\phi}_k I_{i\Delta}^k\right)^2\right]^{1/2}} \Longrightarrow \frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1}{\left(\tilde{\Psi}_3 - \tilde{\Psi}_2^2\right)^{1/2}}.
$$

■

Proof of Theorem 3.2. Under the alternative hypothesis of (10), from Lemma 3.2, the first term

$$
\Upsilon_T^{-1}\begin{bmatrix} T & \sum I_{i\Delta}^1 & \cdots & \sum I_{i\Delta}^K & \sum y_{(i-1)\Delta} \\ \sum I_{i\Delta}^1 & \sum\left(I_{i\Delta}^1\right)^2 & \cdots & \sum I_{i\Delta}^1 I_{i\Delta}^K & \sum I_{i\Delta}^1 y_{(i-1)\Delta} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \sum I_{i\Delta}^K & \sum I_{i\Delta}^K I_{i\Delta}^1 & \cdots & \sum\left(I_{i\Delta}^K\right)^2 & \sum I_{i\Delta}^K y_{(i-1)\Delta} \\ \sum y_{(i-1)\Delta} & \sum y_{(i-1)\Delta} I_{i\Delta}^1 & \cdots & \sum y_{(i-1)\Delta} I_{i\Delta}^K & \sum y_{(i-1)\Delta}^2 \end{bmatrix}\Upsilon_T^{-1} \Rightarrow \begin{bmatrix} 1 & 0 & \cdots & 0 & \sigma N^{1/2}\tilde{\Xi}_2 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \sigma N^{1/2}\tilde{\Xi}_2 & 0 & \cdots & 0 & \sigma^2 N\tilde{\Xi}_3 \end{bmatrix}
$$

and the second term

$$
\Upsilon_T^{-1}\begin{bmatrix} \sum \varepsilon_{i\Delta} \\ \sum I_{i\Delta}^1\varepsilon_{i\Delta} \\ \vdots \\ \sum I_{i\Delta}^K\varepsilon_{i\Delta} \\ \sum y_{(i-1)\Delta}\varepsilon_{i\Delta} \end{bmatrix} \Rightarrow \begin{bmatrix} w_1 \\ \varepsilon_{\tau_1} \\ \vdots \\ \varepsilon_{\tau_K} \\ \sigma N^{1/2}\tilde{\Xi}_4 \end{bmatrix}.
$$

Combing these two terms, we have

$$
\begin{bmatrix} T\left(\tilde{\alpha}-\alpha_0\right) \\ T^{1/2}\left(\tilde{\phi}_1 - \phi_1\right) \\ \vdots \\ T^{1/2}\left(\tilde{\phi}_K - \phi_K\right) \\ T\left(\tilde{\beta}-\beta_0\right) \end{bmatrix} \Rightarrow \sigma N^{1/2}\begin{bmatrix} 1 & 0 & \cdots & 0 & \sigma N^{1/2}\tilde{\Xi}_2 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \sigma N^{1/2}\tilde{\Xi}_2 & 0 & \cdots & 0 & \sigma^2 N\tilde{\Xi}_3 \end{bmatrix}^{-1}\begin{bmatrix} w_1 \\ \varepsilon_{\tau_1} \\ \vdots \\ \varepsilon_{\tau_K} \\ \sigma N^{1/2}\tilde{\Xi}_4 \end{bmatrix} = \begin{bmatrix} \sigma N^{1/2}\frac{\tilde{\Xi}_3 w_1 - \tilde{\Xi}_2\tilde{\Xi}_4}{\tilde{\Xi}_3 - \tilde{\Xi}_2^2} \\ \sigma N^{1/2}\varepsilon_{\tau_1} \\ \vdots \\ \sigma N^{1/2}\varepsilon_{\tau_K} \\ \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1}{\tilde{\Xi}_3 - \tilde{\Xi}_2^2} \end{bmatrix}.
$$

Similarly, we can show that the error variance $\tilde{\sigma}_v^2 = \sum \left(y_{i\Delta} - \tilde{\beta} y_{(i-1)\Delta} - \tilde{\alpha} - \sum_{k=1}^K \tilde{\phi}_k I_{i\Delta}^k \right)^2 = \sigma^2 \Delta \sum \varepsilon_{i\Delta}^2 \left[1 + o_p(1)\right] \to \sigma^2 N$. From Lemma 3.2 and the results that $\tilde{\beta} - \beta_0 = O_p\left(T^{-1}\right)$

$$
\begin{aligned}
DF &= \frac{\left(\tilde{\beta} - 1\right) \left[T \sum_{j=1}^T y_{j\Delta}^2 - \left(\sum_{j=1}^T y_{(j-1)\Delta} \right)^2 \right]^{1/2}}{\left[\sum_{j=1}^T \left(y_{i\Delta} - \tilde{\beta} y_{(i-1)\Delta} - \tilde{\alpha} - \sum_{k=1}^K \tilde{\phi}_k I_{i\Delta}^k \right)^2 \right]^{1/2}} \\
&\Longrightarrow \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1}{\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2 \right)^{1/2}} + c \left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2 \right)^{1/2}.
\end{aligned}
$$

■

Appendix D: the DF Test Statistic in the Presence of Jumps

Proof of Theorem 3.3. The least square estimators of the standardized intercept and the autoregressive coefficients with regression (4), under the DGP of (9), are

$$
\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} - 1 \end{bmatrix} = \begin{bmatrix} T & \sum y_{(j-1)\Delta} \\ \sum y_{(j-1)\Delta} & \sum y_{(j-1)\Delta}^2 \end{bmatrix}^{-1} \left(\begin{bmatrix} \sum_{k=1}^K \phi_k \sum I_{j\Delta}^k \\ \sum_{k=1}^K \phi_k \sum y_{(j-1)\Delta} I_{j\Delta}^k \end{bmatrix} + \sigma \sqrt{\Delta} \begin{bmatrix} \sum \varepsilon_{j\Delta} \\ \sum y_{(j-1)\Delta} \varepsilon_{j\Delta} \end{bmatrix} \right).
$$

Pre-multiplying the above equation by $\Upsilon_{T_n} = diag\left(\sqrt{T}, \sqrt{T}\right)$ leads to

$$
\begin{aligned}
& T^{1/2} \Upsilon_{T_n} \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} - 1 \end{bmatrix} \\
&\quad = \left\{ \Upsilon_{T_n}^{-1} \begin{bmatrix} T & \sum y_{(j-1)\Delta} \\ \sum y_{(j-1)\Delta} & \sum y_{(j-1)\Delta}^2 \end{bmatrix} \Upsilon_{T_n}^{-1} \right\}^{-1} \begin{bmatrix} \sum_{k=1}^K \phi_k + \sigma \sqrt{\Delta} \sum \varepsilon_{j\Delta} \\ \sum_{k=1}^K \phi_k y_{(\tau_k - 1)\Delta} + \sigma \sqrt{\Delta} \sum y_{(j-1)\Delta} \varepsilon_{j\Delta} \end{bmatrix}.
\end{aligned}
$$

The first term

$$
\Upsilon_{T_n}^{-1} \begin{bmatrix} T & \sum y_{(j-1)\Delta} \\ \sum y_{(j-1)\Delta} & \sum y_{(j-1)\Delta}^2 \end{bmatrix} \Upsilon_{T_n}^{-1} = \begin{bmatrix} T & T^{-1} \sum y_{(j-1)\Delta} \\ T^{-1} \sum y_{(j-1)\Delta} & T^{-1} \sum y_{(j-1)\Delta}^2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \sigma N^{1/2} \tilde{\Psi}_2 \\ \sigma N^{1/2} \tilde{\Psi}_2 & \sigma^2 N \tilde{\Psi}_3 \end{bmatrix}.
$$

For the second term,

$$
\begin{aligned}
\begin{bmatrix} \sum_{k=1}^K \phi_k + \sigma \sqrt{\Delta} \sum \varepsilon_{j\Delta} \\ \sum_{k=1}^K \phi_k y_{(\tau_k - 1)\Delta} + \sigma \sqrt{\Delta} \sum y_{(j-1)\Delta} \varepsilon_{j\Delta} \end{bmatrix} &= \begin{bmatrix} \sum_{k=1}^K \phi_k + \sigma N^{1/2} \left(T^{-1/2} \sum \varepsilon_{j\Delta} \right) \\ \sum_{k=1}^K \phi_k y_{(\tau_k - 1)\Delta} + \sigma N^{1/2} \left(T^{-1/2} \sum y_{(j-1)\Delta} \varepsilon_{j\Delta} \right) \end{bmatrix} \\
&\Rightarrow \begin{bmatrix} \sigma N^{1/2} \left(\sum_{k=1}^K \frac{\phi_k}{\sigma N^{1/2}} + w_1 \right) \\ \sigma^2 N \left[\sum_{k=1}^K \frac{\phi_k}{\sigma N^{1/2}} \left(w_{r_k} + \gamma + \Delta_2 \right) + \tilde{\Psi}_4 \right] \end{bmatrix}.
\end{aligned}
$$

Combing these two terms, we have

$$
\begin{bmatrix} T \hat{\alpha} \\ T\left(\hat{\beta} - 1 \right) \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & \sigma N^{1/2} \tilde{\Psi}_2 \\ \sigma N^{1/2} \tilde{\Psi}_2 & \sigma^2 N \tilde{\Psi}_3 \end{bmatrix}^{-1} \begin{bmatrix} \sigma N^{1/2} \left(\sum_{k=1}^K \frac{\phi_k}{\sigma N^{1/2}} + w_1 \right) \\ \sigma^2 N \left[\sum_{k=1}^K \frac{\phi_k}{\sigma N^{1/2}} \left(w_{r_k} + \gamma + \Delta_2 \right) + \tilde{\Psi}_4 \right] \end{bmatrix}
$$

$$= \left[\begin{array}{c} \sigma N^{1/2} \frac{\tilde{\Psi}_3 w_1 - \tilde{\Psi}_2 \tilde{\Psi}_4 + \sum_{k=1}^{K} \frac{\phi_k}{\sigma N^{1/2}} \left[\tilde{\Psi}_3 - \tilde{\Psi}_2 \left(w_{r_k} + \gamma + \Delta_2 \right) \right]}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2} \\ \frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1 + \sum_{k=1}^{K} \frac{\phi_k}{\sigma N^{1/2}} \left(w_{r_k} + \gamma + \Delta_2 - \tilde{\Psi}_2 \right)}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2} \end{array} \right].$$

Furthermore,

$$\begin{aligned}
\hat{\sigma}_v^2 &= \sum \left(y_{j\Delta} - \hat{\beta} y_{(j-1)\Delta} - \hat{\alpha} \right)^2 \\
&= \sum \left[\sigma \sqrt{\Delta} \varepsilon_{j\Delta} - \left(\hat{\beta} - 1 \right) y_{(j-1)\Delta} - \hat{\alpha} + \sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right]^2 \\
&= \sigma^2 \Delta \sum \varepsilon_j^2 + \left(\hat{\beta} - 1 \right)^2 \sum y_{(j-1)\Delta}^2 + \hat{\alpha}^2 + \sum \left(\sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right)^2 \\
&\quad -2\sigma\sqrt{\Delta} \left(\hat{\beta} - 1 \right) \sum y_{(j-1)\Delta} \varepsilon_{j\Delta} - 2\sigma\sqrt{\Delta}\hat{\alpha} \sum \varepsilon_j + 2\sigma\sqrt{\Delta} \sum \left(\sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right) \varepsilon_{j\Delta} \\
&\quad +2 \left(\hat{\beta} - 1 \right) \hat{\alpha} \sum y_{(j-1)\Delta} - 2 \left(\hat{\beta} - 1 \right) \sum y_{(j-1)\Delta} \sum_{k=1}^{K} \phi_k I_{j\Delta}^k - 2\hat{\alpha} \sum \sum_{k=1}^{K} \phi_k I_{j\Delta}^k \\
&= \left[\sigma^2 \Delta \sum \varepsilon_j^2 + \sum \left(\sum_{k=1}^{K} \phi_k I_{j\Delta}^k \right)^2 \right] \{ 1 + o_p(1) \} \\
&\to \sigma^2 N \left(1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N} \right)
\end{aligned}$$

from Lemma 3.1 and the fact that $\tilde{\alpha} = O_p \left(T^{-1} \right)$, and $\tilde{\beta} - 1 = O_p \left(T^{-1} \right)$. The DF statistic is

$$\begin{aligned}
DF &= \frac{\left(\hat{\beta} - 1 \right) \left[T \sum_{j=1}^{T} y_{j\Delta}^2 - \left(\sum_{j=1}^{T} y_{(j-1)\Delta} \right)^2 \right]^{1/2}}{\left[\sum_{j=1}^{T} \left(y_{j\Delta} - \hat{\alpha} - \hat{\beta} y_{(j-1)\Delta} \right)^2 \right]^{1/2}} \\
&\Longrightarrow \left[\frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2} + \frac{\sum_{k=1}^{K} \frac{\phi_k}{\sigma N^{1/2}} \left(w_{r_k} + \gamma + \Delta_2 - \tilde{\Psi}_2 \right)}{\tilde{\Psi}_3 - \tilde{\Psi}_2^2} \right] \left(\frac{\tilde{\Psi}_3 - \tilde{\Psi}_2^2}{1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}} \right)^{1/2} \\
&= \frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1 + \sum_{k=1}^{K} \frac{\phi_k}{\sigma N^{1/2}} \left(w_{r_k} + \gamma + \Delta_2 - \tilde{\Psi}_2 \right)}{\left(1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N} \right)^{1/2} \left(\tilde{\Psi}_3 - \tilde{\Psi}_2^2 \right)^{1/2}}.
\end{aligned}$$

■

Proof of Theorem 3.4. The least square estimators of the standardized intercept and the autoregressive coefficients with regression (4), under the DGP of (10), are

$$\begin{aligned}
\left[\begin{array}{c} \hat{\alpha} - \alpha_0 \\ \hat{\beta} - \beta_0 \end{array} \right] &= \left[\begin{array}{cc} T & \sum y_{(j-1)\Delta} \\ \sum y_{(j-1)\Delta} & \sum y_{(j-1)\Delta}^2 \end{array} \right]^{-1} \\
&\quad \left(\left[\begin{array}{c} \sum_{k=1}^{K} \phi_k \sum I_{j\Delta}^k \\ \sum_{k=1}^{K} \phi_k \sum y_{(j-1)\Delta} I_{j\Delta}^k \end{array} \right] + \lambda_0 \left[\begin{array}{c} \sum \varepsilon_{j\Delta} \\ \sum y_{(j-1)\Delta} \varepsilon_{j\Delta} \end{array} \right] \right).
\end{aligned}$$

Pre-multiplying the above equation by $\Upsilon_{T_n} = diag\left(\sqrt{T}, \sqrt{T}\right)$ leads to

$$
\begin{aligned}
& T^{1/2}\Upsilon_{T_n}\left[\begin{array}{c} \hat{\alpha}-\alpha_0 \\ \hat{\beta}-\beta_0 \end{array}\right] \\
&= \left\{\Upsilon_{T_n}^{-1}\left[\begin{array}{cc} T & \sum y_{(j-1)\Delta} \\ \sum y_{(j-1)\Delta} & \sum y_{(j-1)\Delta}^2 \end{array}\right]\Upsilon_{T_n}^{-1}\right\}^{-1}\left[\begin{array}{c} \sum_{k=1}^{K}\phi_k + \lambda_0 \sum \varepsilon_{j\Delta} \\ \sum_{k=1}^{K}\phi_k y_{(\tau_k-1)\Delta} + \lambda_0 \sum y_{(j-1)\Delta}\varepsilon_{j\Delta} \end{array}\right].
\end{aligned}
$$

The first term

$$
\Upsilon_{T_n}^{-1}\left[\begin{array}{cc} T & \sum y_{(j-1)\Delta} \\ \sum y_{(j-1)\Delta} & \sum y_{(j-1)\Delta}^2 \end{array}\right]\Upsilon_{T_n}^{-1} = \left[\begin{array}{cc} T & T^{-1}\sum y_{(j-1)\Delta} \\ T^{-1}\sum y_{(j-1)\Delta} & T^{-1}\sum y_{(j-1)\Delta}^2 \end{array}\right] \Rightarrow \left[\begin{array}{cc} 1 & \sigma N_2^{1/2}\tilde{\Xi}_2 \\ \sigma N^{1/2}\tilde{\Xi}_2 & \sigma^2 N\tilde{\Xi}_3 \end{array}\right].
$$

For the second term,

$$
\begin{aligned}
\left[\begin{array}{c} \sum_{k=1}^{K}\phi_k + \lambda_0 \sum \varepsilon_{j\Delta} \\ \sum_{k=1}^{K}\phi_k y_{(\tau_k-1)\Delta} + \lambda_0 \sum y_{(j-1)\Delta}\varepsilon_{j\Delta} \end{array}\right] &= \left[\begin{array}{c} \sum_{k=1}^{K}\phi_k + T^{1/2}\lambda_0\left(T^{-1/2}\sum \varepsilon_{j\Delta}\right) \\ \sum_{k=1}^{K}\phi_k y_{(\tau_k-1)\Delta} + T^{1/2}\lambda_0\left(T^{-1/2}\sum y_{(j-1)\Delta}\varepsilon_{j\Delta}\right) \end{array}\right] \\
&\Longrightarrow \left[\begin{array}{c} \sigma N^{1/2}\left(\sum_{k=1}^{K}\varsigma_k + w_1\right) \\ \sigma^2 N\left\{\sum_{k=1}^{K}\varsigma_k\left[\delta\left(1-e^{r_k c}\right) + J_c\left(r_k\right) + e^{r_k c}\gamma + \Delta_4\right] + \tilde{\Xi}_4\right\} \end{array}\right].
\end{aligned}
$$

Combing these two terms, we have

$$
\begin{aligned}
& \left[\begin{array}{c} T\left(\hat{\alpha}-\alpha_0\right) \\ T\left(\hat{\beta}-\beta_0\right) \end{array}\right] \\
\Rightarrow & \left[\begin{array}{cc} 1 & \sigma N_2^{1/2}\tilde{\Xi}_2 \\ \sigma N_2^{1/2}\tilde{\Xi}_2 & \sigma^2 N\tilde{\Xi}_3 \end{array}\right]^{-1}\left[\begin{array}{c} \sigma N^{1/2}\left(\sum_{k=1}^{K}\varsigma_k + w_1\right) \\ \sigma^2 N\left\{\sum_{k=1}^{K}\varsigma_k\left[\delta\left(1-e^{r_k c}\right) + J_c\left(r_k\right) + e^{r_k c}\gamma + \Delta_4\right] + \tilde{\Xi}_4\right\} \end{array}\right] \\
= & \left[\begin{array}{c} \sigma N^{1/2}\frac{\tilde{\Xi}_3 w_1 - \tilde{\Xi}_2\tilde{\Xi}_4 + \sum_{k=1}^{K}\varsigma_k\left[\tilde{\Xi}_3 - \left[\delta\left(1-e^{r_k c}\right) + J_c(r_k) + e^{r_k c}\gamma + \Delta_4\right]\tilde{\Xi}_2\right]}{\tilde{\Xi}_3 - \tilde{\Xi}_2^2} \\ \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1 + \sum_{k=1}^{K}\varsigma_k\left[\delta\left(1-e^{r_k c}\right) + J_c(r_k) + e^{r_k c}\gamma + \Delta_4 - \tilde{\Xi}_2\right]}{\tilde{\Xi}_3 - \tilde{\Xi}_2^2} \end{array}\right].
\end{aligned}
$$

Furthermore,

$$
\begin{aligned}
\hat{\sigma}_v^2 &= \sum\left(y_{j\Delta} - \hat{\beta}y_{(j-1)\Delta} - \hat{\alpha}\right)^2 \\
&= \sum\left[\lambda_0\varepsilon_{j\Delta} - \left(\hat{\beta}-\beta_0\right)y_{(j-1)\Delta} - \left(\hat{\alpha}-\alpha_0\right) + \sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right]^2 \\
&= \lambda_0^2\sum\varepsilon_j^2 + \left(\hat{\beta}-\beta_0\right)^2\sum y_{(j-1)\Delta}^2 + \left(\hat{\alpha}-\alpha_0\right)^2 + \sum\left(\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right)^2 \\
&\quad -2\lambda_0\left(\hat{\beta}-\beta_0\right)\sum y_{(j-1)\Delta}\varepsilon_{j\Delta} - 2\lambda_0\left(\hat{\alpha}-\alpha_0\right)\sum\varepsilon_j + 2\lambda_0\sum\left(\sum_{k=1}^{K}\phi_k I_{j\Delta}^k\right)\varepsilon_{j\Delta} \\
&\quad +2\left(\hat{\beta}-1\right)\left(\hat{\alpha}-\alpha_0\right)\sum y_{(j-1)\Delta} - 2\left(\hat{\beta}-1\right)\sum y_{(j-1)\Delta}\sum_{k=1}^{K}\phi_k I_{j\Delta}^k - 2\left(\hat{\alpha}-\alpha_0\right)\sum\sum_{k=1}^{K}\phi_k I_{j\Delta}^k
\end{aligned}
$$

$$
\begin{aligned}
&= \left[\lambda_0^2 \sum \varepsilon_j^2 + \sum \left(\sum_{k=1}^{K} \phi_k I_{j\Delta}^k\right)^2\right] \{1 + o_p(1)\} \\
&\to \sigma^2 N \left(1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}\right)
\end{aligned}
$$

from Lemma 3.1 and the fact that $\tilde{\alpha} = O_p\left(T^{-1}\right)$, and $\tilde{\beta} - 1 = O_p\left(T^{-1}\right)$. The DF statistic is

$$
\begin{aligned}
DF &= \frac{\left(\hat{\beta} - 1\right)\left[T \sum_{j=1}^{T} y_{j\Delta}^2 - \left(\sum_{j=1}^{T} y_{(j-1)\Delta}\right)^2\right]^{1/2}}{\left[\sum_{j=1}^{T}\left(y_{j\Delta} - \hat{\alpha} - \hat{\beta} y_{(j-1)\Delta}\right)^2\right]^{1/2}} \\
&\Longrightarrow \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1 + \sum_{k=1}^{K} \varsigma_k \left[\delta\left(1 - e^{r_k c}\right) + J_c(r_k) + e^{r_k c}\gamma + \Delta_4 - \tilde{\Xi}_2\right]}{\tilde{\Xi}_3 - \tilde{\Xi}_2^2}\left(\frac{\tilde{\Xi}_3 - \tilde{\Xi}_2^2}{1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}}\right)^{1/2} + c\left(\frac{\tilde{\Xi}_3 - \tilde{\Xi}_2^2}{1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}}\right)^{1/2} \\
&= \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1 + \sum_{k=1}^{K} \varsigma_k \left[\delta\left(1 - e^{r_k c}\right) + J_c(r_k) + e^{r_k c}\gamma + \Delta_4 - \tilde{\Xi}_2\right]}{\left(1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}\right)^{1/2}\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2}} + c\left(\frac{\tilde{\Xi}_3 - \tilde{\Xi}_2^2}{1 + \sum_{k=1}^{K} \frac{\phi_k^2}{\sigma^2 N}}\right)^{1/2}.
\end{aligned}
$$

■

Appendix D: Asymptotics of $DF^{\hat{J}}$

Proof of Theorem 3.5. The null and alternative models can be written in matrix form as follows:

$$
Y = X\theta^0 + \sigma\sqrt{\Delta}\varepsilon \text{ and } Y = X\theta^1 + \lambda_0 \varepsilon,
$$

where $Y = [y_{1\Delta}, y_{2\Delta}, \cdots, y_{T\Delta}]'$, $x_{i\Delta} = \left[1, I_{i\Delta}^1, \ldots, I_{i\Delta}^K, y_{(i-1)\Delta}\right]'$, $X = [x_{1\Delta}, x_{2\Delta}, \ldots, x_{T\Delta}]'$, $\theta^0 = (0, \phi_1, \ldots, \phi_K, 1)$, $\theta^1 = (\alpha_0, \phi_1, \ldots, \phi_K, \beta_0)$, $\varepsilon = [\varepsilon_{1\Delta}, \varepsilon_{2\Delta}, \cdots, \varepsilon_{T\Delta}]'$. The regression model is

$$
Y = \hat{X}\theta + v,
$$

where $\hat{x}_{i\Delta} = \left[1, \hat{I}_{i\Delta}^1, \ldots, \hat{I}_{i\Delta}^{\hat{K}}, y_{(i-1)\Delta}\right]'$, $\hat{X} = [\hat{x}_{1\Delta}, \hat{x}_{2\Delta}, \ldots, \hat{x}_{T\Delta}]'$, $\theta = \left(\alpha, \phi_1, \ldots, \phi_{\hat{K}}, \beta\right)$, and $v = (v_{1\Delta}, v_{2\Delta}, \ldots, v_{T\Delta})'$. Let $\ddot{\theta} = \left(\ddot{\alpha}, \ddot{\phi}_1, \ldots, \ddot{\phi}_{\hat{K}}, \ddot{\beta}\right)'$ be the OLS estimate of θ. We have

$$
\ddot{\theta} = \left(\hat{X}'\hat{X}\right)^{-1} \hat{X}'Y = \left(\hat{X}'\hat{X}\right)^{-1} \hat{X}'X\theta^0 + \sigma\sqrt{\Delta}\left(\hat{X}'\hat{X}\right)^{-1} \hat{X}'\varepsilon
$$

under the null and

$$
\ddot{\theta} = \left(\hat{X}'\hat{X}\right)^{-1} \hat{X}'Y = \left(\hat{X}'\hat{X}\right)^{-1} \hat{X}'X\theta^1 + \lambda_0\left(\hat{X}'\hat{X}\right)^{-1} \hat{X}'\varepsilon
$$

under the alternative. Let $\Upsilon_T = diag\left(\sqrt{T}, 1, \cdots, 1, \sqrt{T}\right)$ and $\Upsilon_T^* = T^{1/2}\Upsilon_T$. We have

$$
\Upsilon_T^{*-1}\hat{X}'\hat{X} = \begin{bmatrix}
1 & T^{-1}\sum \hat{I}_{i\Delta}^1 & \cdots & T^{-1}\sum \hat{I}_{i\Delta}^{\hat{K}} & T^{-1}\sum y_{(i-1)\Delta} \\
T^{-1/2}\sum \hat{I}_{i\Delta}^1 & T^{-1/2}\sum\left(\hat{I}_{i\Delta}^1\right)^2 & \cdots & T^{-1/2}\sum \hat{I}_{i\Delta}^1 \hat{I}_{i\Delta}^{\hat{K}} & T^{-1/2}\sum \hat{I}_{i\Delta}^1 y_{(i-1)\Delta} \\
\vdots & \vdots & \cdots & \vdots & \vdots \\
T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}}\hat{I}_{i\Delta}^1 & \cdots & T^{-1/2}\sum\left(\hat{I}_{i\Delta}^{\hat{K}}\right)^2 & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} y_{(i-1)\Delta} \\
T^{-1}\sum y_{(i-1)\Delta} & T^{-1}\sum y_{(i-1)\Delta}\hat{I}_{i\Delta}^1 & \cdots & T^{-1}\sum y_{(i-1)\Delta}\hat{I}_{i\Delta}^{\hat{K}} & T^{-1}\sum y_{(i-1)\Delta}^2
\end{bmatrix},
$$

$$\Upsilon_T^{*-1}\hat{X}'X = \begin{bmatrix} 1 & T^{-1}\sum I_{i\Delta}^1 & \cdots & T^{-1}\sum I_{i\Delta}^K & T^{-1}\sum y_{(i-1)\Delta} \\ T^{-1/2}\sum \hat{I}_{i\Delta}^1 & T^{-1/2}\sum \hat{I}_{i\Delta}^1 I_{i\Delta}^1 & \cdots & T^{-1/2}\sum \hat{I}_{i\Delta}^1 I_{i\Delta}^K & T^{-1/2}\sum \hat{I}_{i\Delta}^1 y_{(i-1)\Delta} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} I_{i\Delta}^1 & \cdots & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} I_{i\Delta}^K & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} y_{(i-1)\Delta} \\ T^{-1}\sum y_{(i-1)\Delta} & T^{-1}\sum y_{(i-1)\Delta} I_{i\Delta}^1 & \cdots & T^{-1}\sum y_{(i-1)\Delta} I_{i\Delta}^K & T^{-1}\sum y_{(i-1)\Delta}^2 \end{bmatrix},$$

and

$$\Upsilon_T^{-1}\hat{X}'\hat{X}\Upsilon_T^{-1} = \begin{bmatrix} 1 & T^{-1/2}\sum \hat{I}_{i\Delta}^1 & \cdots & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} & T^{-1}\sum y_{(i-1)\Delta} \\ T^{-1/2}\sum \hat{I}_{i\Delta}^1 & \sum \left(\hat{I}_{i\Delta}^1\right)^2 & \cdots & \sum \hat{I}_{i\Delta}^1 \hat{I}_{i\Delta}^{\hat{K}} & T^{-1/2}\sum \hat{I}_{i\Delta}^1 y_{(i-1)\Delta} \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} & \sum \hat{I}_{i\Delta}^{\hat{K}} \hat{I}_{i\Delta}^1 & \cdots & \sum \left(\hat{I}_{i\Delta}^{\hat{K}}\right)^2 & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} y_{(i-1)\Delta} \\ T^{-1}\sum y_{(i-1)\Delta} & T^{-1/2}\sum \hat{I}_{i\Delta}^1 y_{(i-1)\Delta} & \cdots & T^{-1/2}\sum \hat{I}_{i\Delta}^{\hat{K}} y_{(i-1)\Delta} & T^{-1}\sum y_{(i-1)\Delta}^2 \end{bmatrix}.$$

By construction, we have $\sum I_{i\Delta}^k = 1$, $\sum \hat{I}_{i\Delta}^k = 1$, $\sum I_{i\Delta}^s I_{i\Delta}^l = 0$, and $\sum \hat{I}_{i\Delta}^s \hat{I}_{i\Delta}^l = 0$ for any $s, l, k \in \left[1, \hat{K}\right]$ and $s \neq l$. (1) Under the null hypothesis of (9),

$$\Upsilon_T^{*-1}\hat{X}'\hat{X} \implies \begin{bmatrix} 1 & 0 & \cdots & 0 & \sigma N^{1/2}\tilde{\Psi}_2 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \sigma N^{1/2}\tilde{\Psi}_2 & 0 & \cdots & 0 & \sigma^2 N\tilde{\Psi}_3 \end{bmatrix}_{(K+2)\times(K+2)},$$

$$\Upsilon_T^{*-1}\hat{X}'X \implies \begin{bmatrix} 1 & 0 & \cdots & 0 & \sigma N^{1/2}\tilde{\Psi}_2 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \sigma N^{1/2}\tilde{\Psi}_2 & 0 & \cdots & 0 & \sigma^2 N\tilde{\Psi}_3 \end{bmatrix}_{(K+2)\times(K+2)},$$

$$\Upsilon_T^{-1}\hat{X}'\hat{X}\Upsilon_T^{-1} \implies \begin{bmatrix} 1 & 0 & \cdots & 0 & \sigma N^{1/2}\tilde{\Psi}_2 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \\ \sigma N^{1/2}\tilde{\Psi}_2 & 0 & \cdots & 0 & \sigma^2 N\tilde{\Psi}_3 \end{bmatrix}_{(K+2)\times(K+2)}$$

using results from Lemma 3.1. Therefore,

$$\ddot{\theta} = \left(\Upsilon_T^{*-1}\hat{X}'\hat{X}\right)^{-1}\left(\Upsilon_T^{*-1}\hat{X}'X\right)\theta^0 + \sigma\sqrt{\Delta}\left(\hat{X}'\hat{X}\right)^{-1}\hat{X}'\varepsilon \sim \theta^0 + \sigma\sqrt{\Delta}\left(\hat{X}'\hat{X}\right)^{-1}\hat{X}'\varepsilon.$$

Furthermore,

$$T^{1/2}\Upsilon_T^{-1}\sigma\sqrt{\Delta}\hat{X}'\varepsilon = T^{1/2}\sigma\sqrt{\Delta}\begin{bmatrix} T^{-1/2}\sum \varepsilon_{i\Delta} \\ \varepsilon_{\hat{\tau}_1\Delta} \\ \vdots \\ \varepsilon_{\hat{\tau}_k\Delta} \\ T^{-1/2}\sum y_{(i-1)\Delta}\varepsilon_{i\Delta} \end{bmatrix} \implies \sigma N^{1/2}\begin{bmatrix} w_1 \\ \varepsilon_{\tau_1\Delta} \\ \vdots \\ \varepsilon_{\tau_k\Delta} \\ \sigma N^{1/2}\tilde{\Psi}_4 \end{bmatrix}$$

using results from Lemma 3.1 and Remark 3.2. Therefore,

$$T^{1/2}\Upsilon_T\left(\ddot{\theta}-\theta^0\right)\sim T^{1/2}\sigma\sqrt{\Delta}\left(\Upsilon_T^{-1}\hat{X}'\hat{X}\Upsilon_T^{-1}\right)^{-1}\Upsilon_T^{-1}\hat{X}'\varepsilon\Longrightarrow\begin{bmatrix}\sigma N^{1/2}\frac{\tilde{\Psi}_3w_1-\tilde{\Psi}_2\tilde{\Psi}_4}{\tilde{\Psi}_3-\tilde{\Psi}_2^2}\\ \sigma N^{1/2}\varepsilon_{\tau_1}\\ \vdots\\ \sigma N^{1/2}\varepsilon_{\tau_K}\\ \frac{\tilde{\Psi}_4-\tilde{\Psi}_2w_1}{\tilde{\Psi}_3-\tilde{\Psi}_2^2}\end{bmatrix}.$$

(2) Under the alternative of (10)

$$\Upsilon_T^{*-1}\hat{X}'\hat{X}\Longrightarrow\begin{bmatrix}1&0&\cdots&0&\sigma N^{1/2}\tilde{\Xi}_2\\0&1&\cdots&0&0\\\vdots&\vdots&\cdots&\vdots&\vdots\\0&0&\cdots&1&0\\\sigma N^{1/2}\tilde{\Xi}_2&0&\cdots&0&\sigma^2N\tilde{\Xi}_3\end{bmatrix}_{(K+2)\times(K+2)},$$

$$\Upsilon_T^{*-1}\hat{X}'X\Longrightarrow\begin{bmatrix}1&0&\cdots&0&\sigma N^{1/2}\tilde{\Xi}_2\\0&1&\cdots&0&0\\\vdots&\vdots&\cdots&\vdots&\vdots\\0&0&\cdots&1&0\\\sigma N^{1/2}\tilde{\Xi}_2&0&\cdots&0&\sigma^2N\tilde{\Xi}_3\end{bmatrix}_{(K+2)\times(K+2)},$$

$$\Upsilon_T^{-1}\hat{X}'\hat{X}\Upsilon_T^{-1}\Longrightarrow\begin{bmatrix}1&0&\cdots&0&\sigma N^{1/2}\tilde{\Xi}_2\\0&1&\cdots&0&0\\\vdots&\vdots&\cdots&\vdots&\vdots\\0&0&\cdots&1&0\\\sigma N^{1/2}\tilde{\Xi}_2&0&\cdots&0&\sigma^2N\tilde{\Xi}_3\end{bmatrix}_{(K+2)\times(K+2)}$$

using results from Lemma 3.2. Therefore,

$$\ddot{\theta}=\left(\Upsilon_T^{*-1}\hat{X}'\hat{X}\right)^{-1}\left(\Upsilon_T^{*-1}\hat{X}'X\right)\theta^1+\lambda_0\left(\hat{X}'\hat{X}\right)^{-1}\hat{X}'\varepsilon\sim\theta^1+\lambda_0\left(\hat{X}'\hat{X}\right)^{-1}\hat{X}'\varepsilon.$$

Furthermore,

$$T^{1/2}\Upsilon_T^{-1}\lambda_0\hat{X}'\varepsilon=T^{1/2}\lambda_0\begin{bmatrix}T^{-1/2}\sum\varepsilon_{i\Delta}\\ \varepsilon_{\hat{\tau}_1\Delta}\\ \vdots\\ \varepsilon_{\hat{\tau}_k\Delta}\\ T^{-1/2}\sum y_{(i-1)\Delta}\varepsilon_{i\Delta}\end{bmatrix}\Longrightarrow\sigma N^{1/2}\begin{bmatrix}w_1\\ \varepsilon_{\tau_1\Delta}\\ \vdots\\ \varepsilon_{\tau_k\Delta}\\ \sigma N^{1/2}\tilde{\Xi}_4\end{bmatrix}$$

using results from Lemma 3.2 and Remark 3.2. Therefore, we have

$$T^{1/2}\Upsilon_T\left(\ddot{\theta}-\theta^1\right)\sim T^{1/2}\lambda_0\left(\Upsilon_T^{-1}\hat{X}'\hat{X}\Upsilon_T^{-1}\right)^{-1}\Upsilon_T^{-1}\hat{X}'\varepsilon\Longrightarrow\begin{bmatrix}\sigma N^{1/2}\frac{\tilde{\Xi}_3w_1-\tilde{\Xi}_2\tilde{\Xi}_4}{\tilde{\Xi}_3-\tilde{\Xi}_2^2}\\ \sigma N^{1/2}\varepsilon_{\tau_1}\\ \vdots\\ \sigma N^{1/2}\varepsilon_{\tau_K}\\ \frac{\tilde{\Xi}_4-\tilde{\Xi}_2w_1}{\tilde{\Xi}_3-\tilde{\Xi}_2^2}\end{bmatrix}$$

under the alternative.

One can see that the limiting properties of $\ddot{\theta}$ are identical to those of $\tilde{\theta} = \left(\tilde{\alpha}, \tilde{\phi}_1, \ldots, \tilde{\phi}_{\hat{K}}, \tilde{\beta}\right)'$ under both the null and the alternative. The remaining part of the proof is analogous to those in Theorem 3.1 and 3.2. That is,

$$DF^{\hat{J}} \Longrightarrow \frac{\tilde{\Psi}_4 - \tilde{\Psi}_2 w_1}{\left(\tilde{\Psi}_3 - \tilde{\Psi}_2^2\right)^{1/2}}$$

under the null hypothesis and

$$DF^{\hat{J}} \Longrightarrow \frac{\tilde{\Xi}_4 - \tilde{\Xi}_2 w_1}{\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2}} + c\left(\tilde{\Xi}_3 - \tilde{\Xi}_2^2\right)^{1/2}$$

under the alternative. ■