Abstract

We develop and justify methodology to consistently test for long-horizon return predictability based on realized variance. To accomplish this, we propose a parametric transaction-level model for the continuous-time log price process based on a pure-jump point process. The model determines the returns and realized variance at any level of aggregation with properties shown to be consistent with the stylized facts in the empirical finance literature. Under our model, the long-memory parameter propagates unchanged from the transaction-level drift to the calendar-time returns and the realized variance, leading endogenously to a balanced predictive regression equation. We propose an asymptotic framework using power-law aggregation in the predictive regression. Within this framework, we propose a hypothesis test for long-horizon return predictability which is asymptotically correctly sized and consistent.
1 Introduction

Research in return predictability has been an active area for decades. Fama and French (1988), Campbell and Schiller (1987, 1988), Mishkin (1990, 1992), Boudoukh and Richardson (1993) found return predictability based on long-run aggregated financial variables such as the dividend yield, price-dividend ratio, and functions of interest rates. In these studies, a predictive regression with the overlapped aggregated return as the response and the similarly overlapped aggregated or non-aggregated financial variable as the predictor, was implemented. As the aggregation horizon increases, long-run return predictability was found to increase.

It has also been well documented in the literature that, even without aggregation, the strong autocorrelation of the predictors employed in these empirical studies induces bias in the OLS estimator of the predictive coefficient. Stambaugh (1999) considered a predictive regression model with an AR(1) predictor. He obtained an expression for the bias of the OLS estimator of the predictive coefficient. Amihud and Hurvich (2004) also considered an AR(1) predictor and introduced the augmented regression method to reduce the bias and studied a bias-corrected hypothesis test for return predictability. Chen, Deo, and Yi (2013) again considered an AR(1) regressor and proposed the quasi restricted likelihood ratio test (QRLRT) for inference on the predictive coefficient. They proposed a test that maintained the nominal size uniformly as the AR(1) coefficient approaches 1 while delivering higher power than the competing procedures. Boudoukh, Richardson, and Whitelaw (2008) considered a sequence of predictive regressions where overlapped aggregated returns at multiple horizons were regressed on an autocorrelated predictor without aggregation. They showed that the OLS estimators of the predictive coefficients were highly correlated across horizons under the assumption of no return predictability, inflating the size of a joint test that assumes no correlation. They conducted a joint Wald test for the OLS estimators of AR(1) predictors and found much weaker evidence of return predictability.

Furthermore, the overlapping aggregation of an originally-autocorrelated predictor strengthens the degree of autocorrelation and increases the likelihood of spurious findings of long-run predictability of aggregated returns. Valkanov (2003) demonstrated that the ordinary $t$-test in an OLS regression with overlapped aggregated autocorrelated regressors tends to
over-reject the null hypothesis of no return predictability.

Kostakis, Magdalinos, and Stamatogiannis (2018) considered a predictive regression on short-memory, integrated, and local-to-unity predictors. They proved that the long-horizon OLS estimator in a predictive regression is inconsistent when the predictor is integrated or is studied in a local-to-unity framework. Using a procedure that robustified against the unknown persistence of the predictor, they found far weaker evidence of return predictability with increasing aggregation horizon compared with most previous empirical literature.

The literature described above considers regressors that either have short memory, are integrated, or are viewed in a local-to-unity framework. This literature does not consider regressors that are fractionally integrated. Nevertheless, it has been established that predictors related to volatility, such as realized variance or VIX, are indeed fractionally integrated. For example, Andersen et al (2001) found that realized stock return volatility is well described by a long-memory process. Bandi and Perron (2006) found a fractionally-cointegrated relationship between realized and implied volatility, suggesting that both series are long-memory processes. Sizova (2013) considered a long-memory predictor in a predictive regression and proved that under certain assumptions as the level of aggregation increases, the population correlation between future returns and lagged realized variance converges to a constant which is a function of the long-memory parameter of the predictor. When the value of the long-memory parameter is zero, this population correlation converges to zero. Hence her theorem establishes that, under certain assumptions, a short-memory regressor does not lead to long-run return predictability as the level of aggregation increases.

The association between long-run future returns and current and past variance has been extensively studied in the financial economics literature. Bansal and Yaron (2004), Bollerslev, Tauchen and Zhou (2009), and Drechsler and Yaron (2011) considered the impact of long-run economic uncertainty on consumption fluctuations, and thus the effect on the expected long-run returns. They proved that the variance premium, defined as the difference between the variance of returns measured under the risk-neutral and physical probability measures, was a good proxy for the latent economic shocks and was correlated with the expected excess returns. Hence these models imply that the variance premium predicts long-run returns. In an empirical study, Bandi and Perron (2008) found that the long-run excess market return had a stronger correlation with past realized variance than it had with the classical...
dividend yield or the consumption-to-wealth ratio proposed by Lettau and Ludvigson (2001). Hence following Bandi and Perron (2008), we will use realized variance as a predictor in our predictive regressions.

In spite of the long-memory property of the realized variance, Bandi and Perron (2008) did not model market variance as a fractionally integrated process. As they pointed out, if stock returns have short memory, then a predictive regression with a long-memory regressor would be unbalanced. Indeed, rudimentary analysis of stock returns suggests that they have short memory although a weak long-memory component would be hard to detect. Researchers have adopted two frameworks to reconcile the apparently unbalanced nature of this predictive regression. One is to impose assumptions on the functional relationship between the time series of returns and regressors as done by Sizova (2013), who assumed that the predictor is in the domain of attraction of a fractional Brownian motion, and that the error term is additive and has shorter memory than the predictor, so that both sides of the predictive regression equation have long memory with the same memory parameter. The other is to directly model the continuous-time log price process, which then determines the returns and regressor at any level of aggregation. Our work falls into the second framework. Under our model, the long-memory parameter propagates unchanged from the transaction level drift to the calendar-time returns and the realized variance with the same memory parameter, leading endogenously to a balanced predictive regression equation.

In the world of ultra-high-frequency data, the actual transaction-level stock price is naturally viewed as a pure-jump process, that is, a marked point process where the points are the transaction times and the marks are the prices. The stock price observed in continuous time is a step function as opposed to a diffusion process. Engle and Russell (1998) studied the time intervals, i.e. durations, between successive transactions and proposed the ACD (Autoregressive Conditional Duration) model for the durations. Since their seminal work, several studies have been conducted to investigate the propagation of transaction-level properties of pure-jump processes under aggregation to lower-frequency time series in discrete time. Deo, Hsieh, and Hurvich (2010) studied the intertrade durations, counts, (i.e. the number of trade), squared returns, and realized variance of 10 NYSE stocks. They found the presence of long memory in all of theses series. In light of this stylized fact, they proposed the LMSD (Long Memory Stochastic Duration) model for the durations. Deo et al
(2009) provided sufficient conditions for the propagation of the long-memory parameter of durations to the corresponding counts and realized variance.

Cao, Hurvich, and Soulier (2017) studied the effect of drift in pure-jump transaction-level models for asset prices in continuous time, driven by a point process. Under their model, the drift is proportional to the driving point process itself, i.e. the cumulative number of transactions. This link reveals a mechanism by which long memory of the intertrade durations leads to long memory in the returns, with the same memory parameter. Our proposed price model follows Cao, Hurvich, and Soulier’s (2017) framework. The calendar-time return derived by our model has a component that is proportional to the counts. Under this model, both calendar-time returns and realized variance are long-memory processes, which leads to a balanced predictive regression.¹

Our paper makes the following contributions to the existing literature. We propose a parametric transaction-level model for the log price. Our model for the log price implies properties of the calendar-time returns and realized variance that are consistent with the stylized facts. We propose an estimation procedure for the model parameters that is easy to implement. For assessing return predictability, we propose to aggregate $\tilde{H} = \tilde{T}^\kappa$ ($\kappa \in (0, 1)$) calendar-time returns in contrast to Sizova’s choice of $\tilde{H} = \theta \tilde{T}$ ($\theta \in (0, 1)$), where $\tilde{T}$ is the number of available calendar-time returns. Sizova (2013) used the hypothesis test of Valkanov (2003) for long-run return predictability under the linear aggregation framework $\tilde{H} = \theta \tilde{T}$, but did not establish the consistency of this test. Within the power-law framework $\tilde{H} = \tilde{T}^\kappa$, we propose a hypothesis test for long-run return predictability based on the sample correlation between the future aggregated returns and the past realized variance. We establish a central limit theorem for the test statistic under the null hypothesis of no long-run return predictability. Our test is consistent, unlike the one used by Sizova (2013). We also provide simulations on a parametric bootstrap approach to testing for long-run return predictability under the power-law framework, and find that the test has a high power while

¹Note that Bollerslev, Sizova, and Tauchen (2012) also considered generalizing their framework by modeling the volatility of the consumption growth rate as a long-memory process in the continuous-time framework of Comte and Renault (1996). Though their generalized model also leads to a balanced predictive regression (since the components of their variance premium regressor are fractionally-cointegrated and the response is a short memory return series), it would follow from Theorem 3 of Sizova (2013) if its assumptions held that the model of Bollerslev, Sizova, and Tauchen (2012) would not lead to long-run return predictability.
maintaining the nominal size. We discuss the applicability of Sizova’s (2013) assumptions for certain models in Section 5.

The paper is organized as follows. In Section 2, we present our proposed price model. In Section 3, we demonstrate return predictability by evaluating the correlation between the aggregated return and the lagged aggregated variance. In Section 4, we propose a theoretically-based hypothesis test for long-run return predictability and use simulations to evaluate the size and the power of the bootstrap test. We compare our work with Sizova (2013) in Section 5, and conclude the paper in Section 6. In Appendix Section A, we propose formulas for estimating the model parameters. In Section B, we describe a procedure for simulating our model. The proofs of our theorems are provided in Section C.

2 The Model

In this section, we propose our transaction-level model based on a pure-jump point process. We then obtain the corresponding calendar-time return series and its properties.

2.1 Cox Process Driven by Fractional Brownian Motion

We start with the Cox process which is a key ingredient in our model. The points of any point process $N$ consist of the sequence $\{T_i, i \in \mathbb{Z}\}$ (here, the transaction arrival times) such that $T_i < T_{i+1}$ for all $i \in \mathbb{Z}$ and $T_0 \leq 0 < T_1$. For all measurable sets $A \subset \mathbb{R}$, we define the point process $N$ by

$$N(A) = \sum_{i \in \mathbb{Z}} 1_{A}(T_i).$$

Let $\lambda(t), t \in \mathbb{R}$ be a non-negative stochastic intensity function, and define the random measure $\Lambda$ by

$$\Lambda(A) = \int_A \lambda(t)dt,$$  \hspace{1cm} (2.1)

for all measurable sets $A$. A Cox process, $N$, with mean measure $\Lambda$, is defined as follows: conditionally on $\Lambda$, $N$ is a Poisson process with mean measure $\Lambda$. 

We now make further assumptions on the stochastic intensity function which guarantee long memory in returns and volatility when \( d > 0 \), where \( d \) is the long-memory parameter. Let

\[
\lambda(t) = \lambda e^{Z_H(t)}, \quad t \in \mathbb{R}, \quad \lambda > 0,
\]

where \( Z_H \) is a Gaussian stationary process with mean zero, variance one and Hurst index \( 1/2 \leq H < 1 \) (where \( H = \frac{1}{2} + d \)), which implies that the lag-\( r \) autocovariance function \( \gamma_Z \) of \( Z_H \) satisfies

\[
\lim_{r \to \infty} r^{2-2H} \gamma_Z(r) = c_Z > 0, \quad H \in (1/2, 1).
\]

Equation (2.3) implies that the autocovariance of \( Z_H \) is positive for large lags. The unconditional mean of \( N(A) \) is equal to

\[
\mathbb{E}[N(A)] = \mathbb{E}[\mathbb{E}[N(A)|A]] = \lambda e^{1/2} \text{Leb}(A).
\]

An example of such a process \( Z_H \) satisfying (2.3) is fractional Gaussian noise defined as the increment process

\[
Z_H(t) = B_H(ct) - B_H(ct - 1),
\]

where \( B_H(t) \) is fractional Brownian motion and \( c > 0 \) is a time-scaling constant to ensure that our model is invariant to the choice of the time unit. Under Equation (2.5), which we assume for the remainder of the paper, \( Z_H \) has autocovariance function given by

\[
\gamma_Z(r) = \frac{1}{2} \left[ |cr + 1|^{2H} - 2|cr|^{2H} + |cr - 1|^{2H} \right],
\]

\( \forall r \in \mathbb{R}, \gamma_Z(r) \geq 0 \) (see Lemma C.6).

For the short-memory case of \( H = \frac{1}{2} \) (\( d = 0 \)), if \( |r| \geq \frac{1}{c} \), \( \gamma_Z(r) = 0 \); if \( |r| < \frac{1}{c} \), \( \gamma_Z(r) = 1 - c|r| \). (see Lemma C.7). We show in Theorem 3.2 that when \( H = \frac{1}{2} \) there is no long-run predictability of aggregated future returns.

### 2.2 Log Price Model

Let \( N(t) = N((0, t]) \) denote the number of transactions in \((0, t]\). For all \( t \geq 0 \), the log price process is defined as

\[
\log P(t) = \sum_{k=1}^{N(t)} \tilde{e}_k,
\]
where
\[ \tilde{e}_k = \mu + e_k , \] (2.8)
\( \mu > 0 \) is the drift term and \( \{e_k\} \) are the efficient shocks, which are assumed to be i.i.d. random variables with mean zero and variance \( \sigma^2_e \), independent of \( N \). The efficient shock \( e_k \) is assumed to reflect the true value of a stock and thus causes a permanent change to the log price. Under (2.8), Equation (2.7) can be expressed as
\[ \log P(t) = \sum_{k=1}^{N(t)} \{\mu + e_k\} = \mu N(t) + \sum_{k=1}^{N(t)} e_k . \] (2.9)
This together with Equation (2.4) imply that \( \mathbb{E}[\log P(t)] = \mu \mathbb{E}[N(t)] = \mu (\lambda e^{2t}) > 0 \). Hence the expected log price is a linear function of \( t \), which reflects the long-term growth of stock price as observed in the empirical finance literature. For example, the average real annual returns of the Standard and Poor 500 Index over the 90-year period from 1889 to 1978 is about 7% (see Mehra and Prescott(1985)).

Though we will not pursue it in this paper, the model in Equation (2.9) could be generalized to multiple drift terms and point processes, for example
\[ \log P(t) = \sum_{k=1}^{N_1(t)} \{\mu_1 + e_{1,k}\} + \sum_{k=1}^{N_2(t)} \{\mu_1 + e_{2,k}\} = \mu_1 N_1(t) + \mu_2 N_2(t) + \sum_{k=1}^{N_1(t)} e_{1,k} + \sum_{k=1}^{N_2(t)} e_{2,k} , \] (2.10)
where \( N_1(t) \) and \( N_2(t) \) are mutually independent counting processes; \( \{e_{1,k}\} \) and \( \{e_{2,k}\} \) are i.i.d. efficient shocks independent of \( N_1(t) \) and \( N_2(t) \), both with means zero and the same standard deviation \( \sigma^2_e \), \( \mu_1 > 0 \), \( \mu_2 < 0 \), and \( \mu_1 + \mu_2 > 0 \). Here \( N_1(t) \) and \( N_2(t) \) can be thought of representing the numbers of buy and sell transactions. As a result, the observed price reflects the net effect of buy and sell transactions. An active period for \( N_1 \) puts upward pressure on the price, while an active period for \( N_2 \) exerts downward pressure on the price. The assumption \( \mu_1 + \mu_2 > 0 \) reflects the long-term upward trend in the log prices. The positive drift term in Equation (2.9) could represent the overall effect of \( \mu_1 \) and \( \mu_2 \) from (2.10). Thus, we consider model (2.9) as a simplified version of model (2.10). As we show in Lemma C.1 and Corollary C.2, both models exhibit properties of long-term return predictability. Henceforth, unless specified explicitly, we will focus on this simplified model (2.9).
2.3 Return series and its properties

The calendar-time return series \( \{r_t\}_{t=1}^{\infty} \) measured at fixed clock-time intervals of width \( \Delta t \) is \( r_t = \log P(t \Delta t) - \log P((t-1) \Delta t) \). Note that \( \Delta t \), for example, can be one minute, five minutes, 30 minutes, or one day. Here we take \( \Delta t = 1 \) (one day) for notational simplicity. Thus the return series becomes \( \{r_t\}_{t=1}^{\infty} \), which is given by

\[
\begin{align*}
    r_t &= \log P(t) - \log P(t-1) = \mu \Delta N(t) + \sum_{k=N(t-1)+1}^{N(t)} e_k ,
\end{align*}
\]

where \( \Delta N(t) = N(t) - N(t-1) \) is the number of transactions within the interval \((t-1, t]\). We refer to the series \( \{\Delta N(t)\}_{t=1}^{\infty} \) as the counts.

Next we present properties of the calendar-time returns \( \{r_t\} \).

**Lemma 2.1.** If \( d > 0 \), for all \( L \in \mathbb{N} \), the lag-\( L \) autocovariance function of the return process \( \{r_t\} \) is given by

\[
\begin{align*}
    \text{cov}(r_0, r_L) &= \mu^2 \lambda^2 e^1 \int_{s=-1}^{0} \int_{t=-1}^{0} \left\{ e^{\gamma Z(L+t-s)} - 1 \right\} dt ds .
\end{align*}
\]

As \( L \to \infty \),

\[
\begin{align*}
    \text{cov}(r_0, r_L) \sim \mu^2 \lambda^2 e \gamma Z(L) .
\end{align*}
\]

**Corollary 2.2.** The variance of \( r_t \) can be represented as

\[
\begin{align*}
    \text{var}(r_t) &= \mu^2 \lambda^2 e^1 \int_{s=-1}^{0} \int_{t=-1}^{0} \left\{ e^{\gamma Z(s-t)} - 1 \right\} dt ds + \lambda e^{1/2} (\mu^2 + \sigma^2_e) .
\end{align*}
\]

**Theorem 2.3.** Let \( \tilde{\gamma}(L) \) be the lag-\( L \) autocovariance function of the squared returns. If \( d > 0 \) then

\[
\begin{align*}
    \tilde{\gamma}(L) &= \text{cov}(r_t^2, r_{t+L}^2) .
\end{align*}
\]

Then

\[
\begin{align*}
    \lim_{L \to \infty} \frac{\tilde{\gamma}(L)}{\gamma Z(L)} = C ,
\end{align*}
\]

where \( C \) is a positive constant. Hence \( \{r_t^2\} \) is a long-memory process.

Lemma 2.1 shows that the returns \( \{r_t\} \) have memory parameter \( d \). In this paper, we consider returns which can be obtained directly from prices rather than excess returns, which cannot.
The long memory of returns may appear to contradict empirical analysis, but the long memory may be hard to detect when embedded in noise. In Table 1 and Figure 1 we present the estimated long memory parameter as well as the ACF and the log-log periodogram plots of returns simulated from the log price model with $d = 0.35$. The estimated long memory parameters based on different bandwidths are all very close to zero and the lags in the ACF plot appear to be statistically insignificant.  

Theorem 2.3 implies that the realized variance is a long-memory process. Therefore, the calendar-time return and the realized variance derived from our model are long memory processes with the same memory parameter, which leads endogenously to a balanced predictive regression equation.

Figure 1: One realization of 80,000 simulated returns from the log-price model. Model parameters: $d = 0.35$, $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$.

Indeed, excess returns appear to have short memory in spite of the fact that in theory they would have memory parameter equal to one if the risk free rate has a unit root as is found empirically.
Table 1: Estimated long memory parameter $\hat{d}_{GPH}$ (a log-log regression of the periodogram versus frequency) based on one realization of 80000 simulated daily returns. Model parameters: $d=0.35$, $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$.

| $d$ | $m = n^{0.3}$ | $m = n^{0.5}$ | $m = n^{0.7}$ |
|-----|----------------|----------------|----------------|
| $\hat{d}_{GPH}$ | -0.0397 | 0.0555 | 0.0204 |

In Appendices [A] and [B] we provide method-of-moments estimators of the model parameters and a methodology for simulating realizations of the log price model. We show that the proposed estimators are consistent in Theorem [A.2] and we evaluate the performance of these estimators in simulations.

3 Return Predictability

We use the aggregated past realized variance as the regressor to predict the aggregated future return. For concreteness, we assume that $\tilde{t} = 1, 2, \ldots, \tilde{T}$ is measured in months, there are $m$ trading days per month, and $r_t$ is the return for the $t^{th}$ day. The aggregated return over the next $\tilde{H}$ months is given by

$$ R_{\tilde{t}, \tilde{t}+\tilde{H}} = \sum_{t=\tilde{t}m+1}^{(\tilde{t}+\tilde{H})m} r_t, $$

(3.1)

The realized variance over the past $\tilde{H}$ months is given by

$$ RV_{\tilde{t} \tilde{H}, \tilde{t}} = \sum_{t=(\tilde{t}-\tilde{H})m+1}^{\tilde{t}m} r_t^2. $$

(3.2)

To evaluate return predictability over the next $\tilde{H}$ months, one can compute the covariance between the future return and the past realized variance, $\text{cov}(R_{\tilde{t}, \tilde{t}+\tilde{H}}, RV_{\tilde{t} \tilde{H}, \tilde{t}})$ or the correlation

$$ \text{corr}(R_{\tilde{t}, \tilde{t}+\tilde{H}}, RV_{\tilde{t} \tilde{H}, \tilde{t}}) = \frac{\text{cov}(R_{\tilde{t}, \tilde{t}+\tilde{H}}, RV_{\tilde{t} \tilde{H}, \tilde{t}})}{\sqrt{\text{var}(R_{\tilde{t}, \tilde{t}+\tilde{H}})\text{var}(RV_{\tilde{t} \tilde{H}, \tilde{t}})}}. $$

(3.3)
Suppose first that \( d > 0 \). In Lemma C.1, we show that \( \text{cov}(R_{\tilde{t}, \tilde{t} + \tilde{H}}, RV_{\tilde{t} - \tilde{H}, \tilde{t}}) \) can be expressed as

\[
\text{cov}(R_{\tilde{t}, \tilde{t} + \tilde{H}}, RV_{\tilde{t} - \tilde{H}, \tilde{t}}) = \sum_{L=1}^{\tilde{H}m} \text{cov}(r_L, r_0^2)L + \sum_{L=\tilde{H}m+1}^{2\tilde{H}m-1} \left(2\tilde{H}m - L\right) \text{cov}(r_L, r_0^2),
\]

(3.4)

and that \( \text{cov}(r_L, r_0^2) > 0 \), for every positive integer \( L \). Hence \( \text{cov}(R_{\tilde{t}, \tilde{t} + \tilde{H}}, RV_{\tilde{t} - \tilde{H}, \tilde{t}}) \) is positive and there is return predictability. As the horizon \( \tilde{H} \) increases, the correlation between \( R_{\tilde{t}, \tilde{t} + \tilde{H}} \) and \( RV_{\tilde{t} - \tilde{H}, \tilde{t}} \) increases, and it converges to a function of the long-memory parameter \( d \):

**Theorem 3.1.** If \( 0 < d < \frac{1}{2} \), as \( \tilde{H} \to \infty \),

\[
\text{corr}(R_{\tilde{t}, \tilde{t} + \tilde{H}}, RV_{\tilde{t} - \tilde{H}, \tilde{t}}) \to (2^d - 1).
\]

This theorem has the same form as Theorem 3 in Sizova (2013), though our assumptions differ from hers. On the other hand, if \( d = 0 \), i.e. short memory, then there is no long-term return predictability as shown in the following theorem:

**Theorem 3.2.** If \( d = 0 \), as \( \tilde{H} \to \infty \),

\[
\text{corr}(R_{\tilde{t}, \tilde{t} + \tilde{H}}, RV_{\tilde{t} - \tilde{H}, \tilde{t}}) \to 0.
\]

Hence as pointed out by Sizova (2013), it is the presence of the long memory component, strengthened by the aggregation of the returns and the squared returns that causes the long-term return predictability. From now on, we use the notation \( \rho \) to represent \( \text{corr}(R_{\tilde{t}, \tilde{t} + \tilde{H}}, RV_{\tilde{t} - \tilde{H}, \tilde{t}}) \) and \( \hat{\rho} \) to represent the sample correlation between these quantities as estimated by

\[
\hat{\rho}_{\tilde{H}, \tilde{I}} = \frac{1}{T - 2\tilde{H}} \sum_{i=\tilde{H}}^{(\tilde{t} - \tilde{H})} \left(\frac{R_{i, i+\tilde{H}} - \overline{R_{i, i+\tilde{H}}}}{RV_{i - \tilde{H}, i} - \overline{RV_{i - \tilde{H}, i}}}\right) \left(\frac{RV_{i - \tilde{H}, i} - \overline{RV_{i - \tilde{H}, i}}}{RV_{i - \tilde{H}, i} - \overline{RV_{i - \tilde{H}, i}}}\right)
\]

\[
\sqrt{\frac{1}{T - 2\tilde{H}} \sum_{i=\tilde{H}}^{(\tilde{t} - \tilde{H})} \left(R_{i, i+\tilde{H}} - \overline{R_{i, i+\tilde{H}}}ight)^2} \sqrt{\frac{1}{T - 2\tilde{H}} \sum_{i=\tilde{H}}^{(\tilde{t} - \tilde{H})} \left(RV_{i - \tilde{H}, i} - \overline{RV_{i - \tilde{H}, i}}\right)^2}
\]

(3.5)

where \( \overline{R_{i, i+\tilde{H}}} \) and \( \overline{RV_{i - \tilde{H}, i}} \) are the sample means of \( R_{i, i+\tilde{H}} \) and \( RV_{i - \tilde{H}, i} \).
4 Statistical Inference on Return Predictability

Theorem 3.2 establishes that when \( d = 0 \), \( \rho \to 0 \) as \( \tilde{H} \to \infty \) and thus there is no long-run return predictability. Hence we wish to test the hypotheses of no long-run return predictability,

\[
H_0 : \quad \rho = 0 \quad (d = 0) \\
H_a : \quad \rho > 0 \quad (d > 0) 
\]

We consider two different asymptotic frameworks for the degree of aggregation \( \tilde{H} \). The first framework is linear growth, \( \tilde{H} = \theta \tilde{T} \), for \( \theta \in (0, 1) \), as used by Sizova (2013). The second framework is a power law, \( \tilde{H} = \tilde{T}^\kappa \), for \( \kappa \in (0, 1) \). Under the linear aggregation framework, we will consider using \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) as the test statistic without further normalization. Under the power-law framework, we will consider the test statistic \( \sqrt{\tilde{T}^{1-\kappa} \hat{\rho}_{\tilde{H}, \tilde{T}}} \), or \( \sqrt{\tilde{T}^{1-\kappa} \tilde{\rho}_{\tilde{H}, \tilde{T}}} \), where \( \tilde{\rho}_{\tilde{H}, \tilde{T}} \) is a slightly modified version of \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) given by (4.1).

Under the linear growth framework, Sizova (2013) proved that \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) converges in distribution to a functional of fractional Brownian Motion subject to several assumptions (e.g. the dynamics of (14) and (15) as well as Assumptions 1 and 2 of that paper for both \( d = 0 \) and \( d > 0 \)). Under these dynamics and assumptions, the resulting asymptotic distribution \( F_\rho(A^d(\tau), B^d(\tau)) \) for \( d = 0 \) could be used to obtain critical values for a hypothesis test based on \( \hat{\rho}_{\tilde{H}, \tilde{T}} \). Though such a test is asymptotically correctly sized, it is inconsistent as implied by Sizova’s (2013) Theorem 4 since the test statistic converges in distribution under both the null hypothesis and the alternative hypothesis. For our model, Figures 2 and 3 show the sampling distribution of \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) under the linear growth framework for \( \tilde{H} \) based on simulated returns with \( d = 0 \) and \( d > 0 \). These results suggest that \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) converges in distribution when \( d = 0 \) and \( d > 0 \), and thus that the power of a hypothesis test for \( d = 0 \) based on \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) will not go to 1 in this framework. Further support for this conclusion is provided by Tables 2 and 3, where we calculate the variance of \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) for several values of \( \tilde{T} \) under the linear growth framework for our model with \( d = 0 \) and \( d > 0 \). The variance of \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) remains essentially constant. Apparently, the power of a correctly-sized test based on \( \hat{\rho}_{\tilde{H}, \tilde{T}} \) would not approach 1 as \( \tilde{T} \to \infty \) under our model in the linear growth framework. Next we obtain the critical values of the correctly-sized test based on the simulated sampling distribution.
of $\hat{\rho}_{H,\tilde{T}}$ from our model under $d = 0$. Tables 6 and 7 show the size and the power of the hypothesis test based on $\hat{\rho}_{H,\tilde{T}}$ under the linear growth framework with nominal size of 5%. This test is approximately correctly sized but has low power even as the sample size increases, as explained above.

We next consider the power-law framework, which we show is more promising than the linear framework for testing for long-run return predictability. The density plots of $\hat{\rho}_{H,\tilde{T}}$ under the power-law framework shown in Figures 4 and 5 and Tables 4 and 5 indicate that as $\tilde{T}$ increases, the variance of $\hat{\rho}_{H,\tilde{T}}$ decreases. Furthermore, the mean of $\hat{\rho}_{H,\tilde{T}}$ increases with $\tilde{T}$ when $d > 0$, perhaps approaching $2^{2d} - 1$, though this is not clear based on the sample sizes considered. This suggests the possibility to construct a consistent test for long-run return predictability based on a rescaled version of $\hat{\rho}_{H,\tilde{T}}$. To determine an appropriate rescaling factor, we study the behavior of the variance of $\hat{\rho}_{H,\tilde{T}}$ as $\tilde{T}$ increases under the null hypothesis, $d = 0$. To study the convergence rate of the variance of $\hat{\rho}_{H,\tilde{T}}$, we created a scatter plot for the logged variance of $\hat{\rho}_{H,\tilde{T}}$ and log $\tilde{T}$ presented in Figure 8. For each value of $\kappa$, the scatter plot appears to be linear. To confirm this, we regress the logged variance of $\hat{\rho}_{H,\tilde{T}}$ on log $\tilde{T}$ and present the fitted equations in Table 11. The fitted slope coefficients are very close to the values of $\kappa - 1$. Hence we conjecture that the variance of $\hat{\rho}_{H,\tilde{T}}$ is proportional to $\tilde{T}^{\kappa - 1}$. This leads to the rescaled test statistic $\sqrt{\tilde{T}^{1-\kappa}}\hat{\rho}_{H,\tilde{T}}$.

Density plots in Figure 6 as well as Shapiro-Wilk normality test results in Table 12 for $d = 0$ indicate that $\sqrt{\tilde{T}^{1-\kappa}}\hat{\rho}_{H,\tilde{T}}$ may be asymptotically normal under the null hypothesis. Therefore, from now on we will focus on the power-law aggregation framework. We will also restrict attention to hypothesis testing under our model. We consider two options for hypothesis testing under the power law framework: (i) bootstrap method and (ii) asymptotic method.

In the bootstrap approach, given one realization of log $P(t)$, $t \in (0, T]$, we compute the observed value of $\sqrt{\tilde{T}^{1-\kappa}}\hat{\rho}_{H,\tilde{T}}$, $\sqrt{\tilde{T}^{1-\kappa}\hat{\rho}_{H,\tilde{T}}^{obs}}$, and estimate the model parameters using the formulas in Section A. We then set $d = 0$ and use the other estimated parameters to simulate

---

1Here we do not consider using the critical values obtained from the asymptotic null distribution in Sizova’s (2013) Theorem 4 for the test. This is because our return model implies an endogenous relationship between the return and the realized variance, whereas Sizova’s (2013) Theorem 4 assumes dynamic (as in her Eq.(15)) under which the realized variance is an exogenous variable.
1000 replications of $\sqrt{T^{1-\kappa}\hat{\rho}_{H,\tilde{t}}}$, We use the 95th percentile among the 1000 replications, $\sqrt{T^{1-\kappa}\hat{\rho}_{95\%}},$ as the critical value of the test. Thus, we compare $\sqrt{T^{1-\kappa}\hat{\rho}_{95\%}}$ with $\hat{\rho}_{95\%}$ and reject the null hypothesis if $\sqrt{T^{1-\kappa}\hat{\rho}_{95\%}} > \sqrt{T^{1-\kappa}\hat{\rho}_{95\%}}$.

Davidson and MacKinnon (1999, 2006) showed that subject to some assumptions, if the model parameter estimator is consistent under the null hypothesis, the rejection rate of a parametric bootstrap test converges to the nominal size, where the convergence rate depends on the order of convergence of the consistent estimator. In Theorem A.2, we prove that our model estimators are consistent under $d = 0$. We find empirically that the rejection rates of our bootstrap hypothesis test (reported in Table 10) are almost never significantly different from the nominal size. Furthermore, the power (see Table 9) increases with the sample size and the power is highest for smaller values of $\kappa$. Nevertheless, we do not have a theoretical justification for the bootstrap hypothesis test in our framework.

Next, we develop an asymptotically correctly-sized consistent test based on $\sqrt{T^{1-\kappa}\hat{\rho}_{H,\tilde{t}}},$ where $\hat{\rho}_{H,\tilde{t}}$ is a slightly modified version of $\hat{\rho}_{H,\tilde{t}}$ (defined in (4.1)). When $d = 0$, our return model implies that the daily return $r_t$ is $h$-dependent, where the lag $h$ is determined by the time-scaling constant $c$ (See (C.50) for the definition of $h$). To take advantage of this fact (see Remark 4.3 below), we propose a modified sample correlation coefficient $\hat{\rho}_{H,\tilde{t}},$ where in the aggregation of returns, we skip the first $h$ daily returns from the first day of month $\tilde{t}$ and aggregate the remaining daily returns over the $H$ months to construct the modified aggregated return $\check{R}_{\tilde{t},\tilde{t}+H}$ and then compute the sample correlation between $\check{R}_{\tilde{t},\tilde{t}+H}$ and $RV_{\tilde{t},\tilde{t}+H}$.

**Theorem 4.1.** Define the modified sample correlation coefficient

$$\hat{\rho}_{H,\tilde{t}} = \frac{1}{T - 2H} \sum_{i=H+1}^{T-H} \left( \check{R}_{\tilde{t},\tilde{t}+H} - \overline{R_H} \right) \left( RV_{\tilde{t}-H,\tilde{t}} - \overline{RV_H} \right)$$

(4.1)

where

$$\check{R}_{\tilde{t},\tilde{t}+H} = \sum_{t=(i+H)m}^{(i+H)m+h+1} r_t$$

(4.2)
and
\[
\bar{R}_{\tilde{H}} = \frac{1}{T - 2\tilde{H}} \sum_{t = H+1}^{T-H} \tilde{R}_{i,i+\tilde{H}}. \quad (4.3)
\]

If \(d = 0\) and \(\tilde{H} = \tilde{T}^{\kappa}\), for \(\kappa \in (0, 1/3)\),
\[
\sqrt{T^{1-\kappa}} \hat{\rho}_{\tilde{H},\tilde{T}} \xrightarrow{(d)} N \left(0, \frac{S^2}{A_1A_2} \right),
\quad (4.4)
\]
where \(A_1 = \text{var}(r_0) + 2 \sum_{k=1}^{h} \text{cov}(r_0, r_k)\), \(A_2 = \text{var}(r_0^2) + 2 \sum_{k=1}^{h} \text{cov}(r_0^2, r_k^2)\), and \(S^2 = \frac{2}{3} \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2)\).

**Remark 4.2.** \(\text{cov}(r_0, r_k), \text{cov}(r_0^2, r_k^2), \text{var}(r_0), \text{and var}(r_0^2)\) can be evaluated by (2.12), (C.32), (A.4), and (C.61) under \(d = 0\).

**Remark 4.3.** The reason we do not attempt to establish a CLT for \(\hat{\rho}_{\tilde{H},\tilde{T}}\) as follows: For any \(d\), we have shown that \(\text{cov}(R_{i,i+\tilde{H}}, RV_{i-\tilde{H},i}) > 0\) (see Lemma C.1). When \(d = 0\), \(\text{cov}(R_{i,i+\tilde{H}}, RV_{i-\tilde{H},i}) = \sum_{k=1}^{h} \text{cov}(r_k, r_0^2)k\) (See (C.51)). This suggests that for \(\kappa \in (0, 1/3)\), \(\sqrt{T^{1-\kappa}} \hat{\rho}_{\tilde{H},\tilde{T}}\) diverges. The modified version \(\bar{R}_{i,i+\tilde{H}}\) is independent of \(RV_{i-\tilde{H},i}\) and thus \(\text{cov}(\bar{R}_{i,i+\tilde{H}}, RV_{i-\tilde{H},i}) = 0\).

Table 2: Variance of 1000 realizations of \(\hat{\rho}_{\tilde{H},\tilde{T}}\) using \(\tilde{H} = \theta T\) when \(d = 0\), \(\mu = 0.000001419188\), \(\lambda = 128.2085\), \(c = 1\), and \(\sigma_e = 0.0007289\).

| \(\theta\)   | Variance | \(\theta\)   | Variance | \(\theta\)   | Variance | \(\theta\)   | Variance | \(\theta\)   | Variance | \(\tilde{T}\) |
|------------|---------|------------|---------|------------|---------|------------|---------|------------|---------|------------|
| 0.0125     | 0.01792 | 0.04084    | 0.10074 | 0.22477    | 131     | 0.00785    | 0.01933 | 0.03776    | 0.09905 | 0.21818    | 262     |
| 0.00785    | 0.01732 | 0.03561    | 0.10008 | 0.21772    | 2097    | 0.00837    | 0.03561 | 0.10008    | 0.21772 | 2097       | 4194    |

16
Table 3: Variance of 1000 realizations of $\hat{\rho}_{H,T}$ using $\tilde{H} = \theta T$ when $d = 0.3545$, $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$.

| $\theta = 0.0125$ | $\theta = 0.025$ | $\theta = 0.05$ | $\theta = 0.125$ | $\theta = 0.25$ | $\tilde{T}$ |
|------------------|------------------|------------------|------------------|------------------|------------|
| 0.01488          | 0.02223          | 0.04795          | 0.10958          | 0.26234          | 131        |
| 0.01163          | 0.02402          | 0.04267          | 0.12113          | 0.27342          | 262        |
| 0.01416          | 0.02460          | 0.04477          | 0.11967          | 0.25633          | 524        |
| 0.01203          | 0.02347          | 0.04523          | 0.12059          | 0.27081          | 1048       |
| 0.01217          | 0.02410          | 0.04785          | 0.11544          | 0.26542          | 2097       |
| 0.01223          | 0.02504          | 0.04867          | 0.12412          | 0.26000          | 4194       |

Table 4: Variance of 1000 realizations of $\hat{\rho}_{H,T}$ using $\tilde{H} = T^\kappa$ when $d = 0$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$.

| $\kappa = 0.1$ | $\kappa = 0.3$ | $\kappa = 0.5$ | $\kappa = 0.7$ | $\tilde{T}$ |
|----------------|----------------|----------------|----------------|------------|
| 0.01259        | 0.02289        | 0.06194        | 0.21163        | 131        |
| 0.00591        | 0.01330        | 0.04637        | 0.15999        | 262        |
| 0.00271        | 0.00959        | 0.03061        | 0.11927        | 524        |
| 0.00146        | 0.00517        | 0.02218        | 0.09242        | 1048       |
| 0.00068        | 0.00313        | 0.01508        | 0.07841        | 2097       |
| 0.00035        | 0.00188        | 0.01026        | 0.06406        | 4194       |
Table 5: Variance of 1000 realizations of \( \hat{\rho}_{H,T} \) using \( \tilde{H} = T^\kappa \) when \( d = 0.3545 \). Model parameters: \( \mu = 0.00001419188 \), \( \lambda = 128.2085 \), \( c = 1 \), and \( \sigma_e = 0.0007289 \).

| \( \kappa = 0.1 \) | \( \kappa = 0.3 \) | \( \kappa = 0.5 \) | \( \kappa = 0.7 \) | \( \hat{T} \) |
|-------------------|-------------------|-------------------|-------------------|--------|
| 0.01488           | 0.02904           | 0.07309           | 0.23814           | 131    |
| 0.00859           | 0.01773           | 0.05195           | 0.19369           | 262    |
| 0.00446           | 0.01416           | 0.04036           | 0.14804           | 524    |
| 0.00242           | 0.00776           | 0.02823           | 0.11967           | 1048   |
| 0.00120           | 0.00491           | 0.02129           | 0.09154           | 2097   |
| 0.00068           | 0.00321           | 0.01529           | 0.08061           | 4194   |

Table 6: Size of Sizova’s (2013) return predictability hypothesis test: \( H_0 : d = 0 \) vs. \( H_a : d > 0 \) using \( \tilde{H} = \theta \hat{T} \). Critical values are obtained from 1000 replications of simulated \( \hat{\rho}_{H,T} \) from our return model.

Model parameters: \( \mu = 0.00001419188 \), \( \lambda = 128.2085 \), \( c = 1 \), and \( \sigma_e = 0.0007289 \). The symbol * indicates rejection of size = 5%.

| \( \theta = 0.0125 \) | \( \theta = 0.025 \) | \( \theta = 0.05 \) | \( \theta = 0.125 \) | \( \theta = 0.25 \) | \( \hat{T} \) |
|-----------------------|-----------------------|-----------------------|-----------------------|-----------------------|--------|
| 0.054                 | 0.058                 | 0.063                 | 0.044                 | 0.055                 | 131    |
| 0.055                 | 0.045                 | 0.052                 | 0.051                 | 0.063                 | 262    |
| 0.054                 | 0.053                 | 0.035*                | 0.056                 | 0.034*                | 524    |
| 0.057                 | 0.056                 | 0.062                 | 0.046                 | 0.036*                | 1048   |
| 0.065                 | 0.053                 | 0.059                 | 0.049                 | 0.035*                | 2097   |
| 0.053                 | 0.090*                | 0.067*                | 0.060                 | 0.044                 | 4194   |
Table 7: Power of return predictability hypothesis test: $H_0 : d = 0$ vs. $H_a : d > 0$ using $\tilde{H} = \theta \tilde{T}$. The power is calculated as the percentage of cases that reject the null hypothesis among the 1000 realizations of return process simulated under $d = 0.3545$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$.

| $\theta = 0.0125$ | $\theta = 0.025$ | $\theta = 0.05$ | $\theta = 0.125$ | $\theta = 0.25$ | $\tilde{T}$ |
|-------------------|-------------------|-------------------|-------------------|-------------------|------------|
| 0.114             | 0.117             | 0.093             | 0.082             | 0.072             | 131        |
| 0.165             | 0.131             | 0.085             | 0.076             | 0.097             | 262        |
| 0.197             | 0.132             | 0.109             | 0.085             | 0.060             | 524        |
| 0.236             | 0.189             | 0.135             | 0.083             | 0.072             | 1048       |
| 0.319             | 0.257             | 0.185             | 0.070             | 0.065             | 2097       |
| 0.394             | 0.372             | 0.217             | 0.088             | 0.048             | 4194       |

Table 8: Size of bootstrap return predictability hypothesis test: $H_0 : d = 0$ vs. $H_a : d > 0$ using $\tilde{H} = \tilde{T}^\kappa$. The size is calculated as the percentage of cases that reject the null hypothesis among the 500 realizations of return process simulated under $d = 0$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$. The symbol $*$ indicates rejection of size $= 5\%$.

| $\kappa = 0.1$ | $\kappa = 0.3$ | $\kappa = 0.5$ | $\kappa = 0.7$ | $\tilde{T}$ |
|----------------|----------------|----------------|----------------|------------|
| 0.055          | 0.062          | 0.057          | 0.054          | 131        |
| 0.042          | 0.041          | 0.042          | 0.055          | 262        |
| 0.051          | 0.051          | 0.066*         | 0.057          | 524        |
| 0.058          | 0.054          | 0.05           | 0.052          | 1048       |
| 0.041          | 0.053          | 0.046          | 0.055          | 2097       |
| 0.051          | 0.068*         | 0.049          | 0.049          | 4194       |
Table 9: Power of bootstrap return predictability hypothesis test: $H_0: d = 0$ vs. $H_a: d > 0$ using $\hat{H} = \hat{T}^\kappa$. The power is calculated as the percentage of cases that reject the null hypothesis among the 1000 realizations of return process simulated under $d = 0.3545$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$.

| $\kappa$ = 0.1 | $\kappa$ = 0.3 | $\kappa$ = 0.5 | $\kappa$ = 0.7 | $\hat{T}$ |
|----------------|----------------|----------------|----------------|----------|
| 0.124          | 0.097          | 0.074          | 0.074          | 131      |
| 0.175          | 0.138          | 0.113          | 0.063          | 262      |
| 0.233          | 0.185          | 0.133          | 0.079          | 524      |
| 0.356          | 0.267          | 0.168          | 0.072          | 1048     |
| 0.518          | 0.407          | 0.235          | 0.108          | 2097     |
| 0.769          | 0.630          | 0.367          | 0.166          | 4194     |

Table 10: Rejection rate from the bootstrap hypothesis test. * indicates rejection of the nominal size 5%.

| $\kappa$ |
|----------|
| 0.1      |
| 0.3      |
| 0.5      |
| 0.7      |
| $\hat{T}$ |
| 0.041    |
| 0.053    |
| 0.046    |
| 0.055    |
| 2000     |
| 0.051    |
| 0.068*   |
| 0.049    |
| 0.049    |
| 4000     |

Table 11: Estimated intercept $\hat{\beta}_0$ and slope $\hat{\beta}_1$ in regression of $\log \hat{\text{var}}(\hat{\rho}_{R,T})$ on $\log \hat{T}$ with $\hat{H} = \hat{T}^\kappa$.

| $\kappa$ |
|----------|
| 0.1      |
| 0.3      |
| 0.5      |
| 0.7      |
| $\hat{\beta}_0$ |
| 0.025    |
| -0.683   |
| -0.501   |
| -0.110   |
| $\hat{\beta}_1$ |
| -0.960   |
| -0.667   |
| -0.486   |
| -0.322   |
| $\kappa - 1$ |
| -0.9     |
| -0.7     |
| -0.5     |
| -0.3     |
Table 12: Shapiro-Wilk normality test on the distribution of $\hat{\rho}_{\tilde{H},\tilde{T}}$ under the linear growth and power-law asymptotic frameworks. The numbers shown in the table are $p$-values. The symbol * indicates rejection of null hypothesis of normality. Model parameters: $d = 0$, $\mu = 0.00001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$.

| $\tilde{H} = \tilde{T}^\kappa$ | $\tilde{H} = \theta \tilde{T}$ |       |       |       |       |       |       |       |
|------------------------------|-------------------------------|-------|-------|-------|-------|-------|-------|-------|
| $\kappa = 0.1$ | $\kappa = 0.3$ | $\kappa = 0.5$ | $\kappa = 0.7$ | $\theta = 0.0125$ | $\theta = 0.025$ | $\theta = 0.05$ | $\theta = 0.125$ | $\theta = 0.25$ | $\tilde{T}$ |
| 0.489 | 0.708 | 0.058 | $< 0.05^*$ | 0.489 | 0.941 | 0.047* | $< 0.05^*$ | $< 0.05^*$ | 131 |
| 0.360 | 0.667 | 0.008* | $< 0.05^*$ | 0.344 | 0.772 | 0.031* | $< 0.05^*$ | $< 0.05^*$ | 262 |
| 0.964 | 0.457 | 0.183 | $< 0.05^*$ | 0.457 | 0.195 | 0.209 | $< 0.05^*$ | $< 0.05^*$ | 525 |
| 0.852 | 0.597 | 0.537 | 0.008* | 0.606 | 0.454 | 0.943 | $< 0.05^*$ | $< 0.05^*$ | 1048 |
| 0.083 | 0.032* | 0.085 | $< 0.05^*$ | 0.310 | 0.108 | 0.296 | $< 0.05^*$ | $< 0.05^*$ | 2097 |
| 0.839 | 0.346 | 0.902 | 0.005* | 0.949 | 0.624 | 0.184 | $< 0.05^*$ | $< 0.05^*$ | 4194 |
Figure 2: Sampling distribution of 1000 realizations of $\hat{\rho}_{\tilde{H},T}$ using $\tilde{H} = \theta T$ when $d = 0$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$. 
Figure 3: Sampling distribution of 1000 realizations of $\hat{\rho}_{\tilde{H},\tilde{T}}$ using $\tilde{H} = \theta T$ when $d = 0.3545$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$. 
Figure 4: Sampling distribution of 1000 realizations of \( \hat{\rho \tilde{H}} \) using \( \tilde{H} = T^\kappa \), \( \kappa = 0.1, 0.3, 0.5, 0.7 \) when \( d = 0 \). Model parameters: \( \mu = 0.00000141988 \), \( \lambda = 128.2085 \), \( c = 1 \), and \( \sigma_e = 0.0007289 \).
Figure 5: Sampling distribution of 1000 realizations of $\hat{\rho}$ when $\tilde{H} = \tilde{T}^{0.1}, \tilde{H} = \tilde{T}^{0.3}, \tilde{H} = \tilde{T}^{0.5}, \tilde{H} = \tilde{T}^{0.7}$ when $d = 0.3545$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$. 
Figure 6: Sampling distribution of 1000 realizations of \( \sqrt{T^{1-\kappa}\hat{\rho}_{R,\bar{t}}} \) using \( \tilde{H} = T^\kappa \), \( \kappa = 0.1, 0.3, 0.5, 0.7 \) when \( d = 0 \). Model parameters: \( \mu = 0.000001419188 \), \( \lambda = 128.2085 \), \( c = 1 \), and \( \sigma_e = 0.0007289 \).
(a) $\tilde{H} = \bar{T}^{0.1}$

(b) $\tilde{H} = \bar{T}^{0.3}$

(c) $\tilde{H} = \bar{T}^{0.5}$

(d) $\tilde{H} = \bar{T}^{0.7}$

Figure 7: Sampling distribution of 1000 realizations of $\sqrt{T^{1-\kappa}\hat{\rho}_{R,\bar{T}}}$ using $\tilde{H} = \bar{T}^{\kappa}$, $\kappa = 0.1, 0.3, 0.5, 0.7$ when $d = 0.3545$. Model parameters: $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$. 
Comparison with Sizova (2013)

Since our paper is not in a local-to-zero framework, we consider a test based on $\hat{\rho}_{\tilde{H},\tilde{T}}$ with critical value obtained from Theorem 4 of Sizova (2013) under the assumption $\tilde{H}/\tilde{T} \to \theta \in (0,1)$. This theorem is based on a further assumption (Assumption 2) that a suitably normalized version of the regressor converges weakly to a fractional Brownian motion. Sizova presented several examples under which Assumptions 1 and 2 in that paper hold. The example that is closest to the context of this paper is Example 3, in which the continuous-time long-memory stochastic volatility model of Comte and Renault (1998) is assumed for the price. In this example, the regressor is taken to be the integrated latent variance, $\sigma^2(t) = \int_0^t \sigma^2(u) du$, where $\sigma^2(u)$ is an unobserved latent variance, assumed to obey a stationary long-memory Ornstein-Uhlenbeck model. Sizova pointed out that for this regressor, Assumption 2 would hold, by an argument similar to the proof of the main result of Taqqu (1976).
comparison with Example 3 of Sizova (2013), we point out that the regressor in the example could not be used in practice even if the model assumed there held since \( \sigma^2(t) \) is a latent process. Clearly, one could try using the realized variance to estimate \( \sigma^2(t) \), and indeed Comte and Renault (1998, Proposition 5.1, Page 305) provide a theoretical result on the \( L^2 \) convergence and \( L^1 \) convergence of a suitably normalized version of realized variance to \( \sigma^2(t) \). However, neither Sizova (2013) nor Comte and Renault (1998) has claimed or proved that Assumption 2 of Sizova (2013) would hold for some function of the realized variance. Furthermore, a practical issue that would arise in trying to estimate \( \sigma^2(t) \) in the model of Comte and Renault (1998) is that realized variance is not a useful quantity when based on very short time increments because the squared returns are mostly zero. This is often attributed to microstructure noise in the literature, but in our viewpoint the problem is deeper, and is most fundamentally attributable to the step-function nature of the log price process.

6 Conclusion

Much of the literature on long-horizon return predictability finds that the evidence for predictability improves as the aggregation horizon increases. Many of these studies, however, were subject to spurious findings due to the amplified correlations in the predictors and the returns caused by the overlapping aggregation.

Our work provides a theoretical framework for exploring the correlation between the long-horizon returns and the realized variance. We propose a parametric transaction-level model for the continuous-time log price based on a pure-jump process. The model determines returns and realized variance at any level of aggregation and reveals a channel through which the statistical properties of the transaction-level data are propagated to induce long-horizon return predictability. Specifically, the memory parameter of returns at the transaction level propagates unchanged to the returns and realized variance at any calendar-time frequency, which leads to a balanced predictive regression. We show that if the long-memory parameter is zero, then there is no long-horizon return predictability. This extends a result shown previously by Sizova (2013) based on discrete-time assumptions. In addition, we provide
consistent estimators for the model parameters. To assess return predictability, we propose a hypothesis test based on a power-law aggregation of the returns and realized variance. We demonstrate that the proposed test is asymptotically correctly-sized and is consistent, whereas the related test of Sizova (2013), which used a linear aggregation framework, is inconsistent.

Appendix

A Model Parameter Estimation

In this section, we derive formulas for estimating the model parameters, \( \mu, \lambda, \sigma_e^2, c, \) and \( d \) based on \( \log P(t), t \in (0, T] \). We use moments of the calendar-time returns \( \{r_t\}_{t=1}^T \) and the counts \( \{\Delta N(t)\}_{t=1}^T \) to derive formulas for estimating model parameters. We establish the consistency of the proposed estimators in Theorem A.2.

Moments of returns and counts

We first obtain certain moments of returns and counts. The expectations of \( r_t \) and \( \Delta N(t) \) can be expressed as

\[
\mathbb{E}[r_t] = \mu \mathbb{E}[\Delta N(t)] \tag{A.1}
\]

where

\[
\mathbb{E}[\Delta N(t)] = \lambda \int_{t-1}^{t} \mathbb{E}[e^{Z_H(s)}]ds = \lambda e^{\frac{1}{2}}. \tag{A.2}
\]

Hence

\[
\mathbb{E}[r_t] = \mu \lambda e^{\frac{1}{2}}. \tag{A.3}
\]

From Corollary 2.2 when \( d = 0 \), Equation (2.13) becomes

\[
\text{var}(r_t) = \mu^2 \lambda^2 e + \lambda e^{\frac{1}{2}} \left( \mu^2 + \sigma_e^2 \right). \tag{A.4}
\]
Lemma A.1. The autocovariance function and the variance of the counts $\Delta N(t)$ are equal to

$$
cov(\Delta N(L), \Delta N(0)) = \lambda^2 \int_{-1}^{0} \int_{-1}^{0} e^{\gamma(Z(L+s-t))} ds dt, \ (L = 1, 2, \cdots) \quad (A.5)
$$

and

$$
var(\Delta N(t)) = \lambda^2 \int_{-1}^{0} \int_{-1}^{0} e^{\gamma(Z(s-t))} ds dt + \lambda e^{\frac{1}{2}}, \quad (A.6)
$$

where

$$
\gamma(Z(L + s - t)) = \frac{1}{2} [c(L + s - t) + 1|2^H - 2|c(L + s - t)|2^H + |c(L + s - t) - 1|2^H] \quad (A.7)
$$

Moreover, when $d = 0$,

$$
var(\Delta N(t)) = \lambda^2 e + \lambda e^{\frac{1}{2}}. \quad (A.8)
$$

Estimation of $\mu$

Let $\bar{r}_t = \frac{1}{T}(r_1 + r_2 + \cdots + r_T)$ and $\overline{\Delta N(t)} = \frac{1}{T}(\Delta N(1) + \Delta N(2) + \cdots + \Delta N(T))$. From Equation (A.1), the drift term can be estimated by

$$
\hat{\mu} = \frac{\bar{r}_t}{\overline{\Delta N(t)}}. \quad (A.9)
$$

Estimation of $\lambda$

Using Equation (A.2), we can estimate $\lambda$ by

$$
\hat{\lambda} = \frac{\overline{\Delta N(t)}}{e^{\frac{1}{2}}}. \quad (A.10)
$$

Estimation of $c$ and $d$

Let $\hat{\gamma}_{\Delta N}(1)$ and $\hat{\gamma}_{\Delta N}(2)$ be the sample lag-one and lag-two autocovariances of the counts. Substituting $\hat{\lambda}$ into Equation (A.5), we solve a system of nonlinear equations to estimate $c$.
and $d$:

\[
\hat{\lambda}^2 \int_{-1}^{0} \int_{-1}^{0} e^{[e^{\gamma Z (s-t)} - 1]} \, ds \, dt - \hat{\gamma}_\Delta N(1) = 0 \quad \text{(A.11)}
\]

\[
\hat{\lambda}^2 \int_{-1}^{0} \int_{-1}^{0} e^{[e^{\gamma Z (s-t)} - 1]} \, ds \, dt - \hat{\gamma}_\Delta N(2) = 0 . \quad \text{(A.12)}
\]

**Estimation of $\sigma_e^2$**

From Equation (A.6), we can express

\[
\lambda e \int_{-1}^{0} \int_{-1}^{0} \{e^{\gamma Z (s-t)} - 1\} \, ds \, dt = \text{var}(\Delta N(t)) - \lambda e^{\frac{1}{2}} \quad \text{(A.13)}
\]

Substituting Equation A.13 into Equation (2.13), Equation (2.13) becomes

\[
\text{var}(r_t) = \mu^2 \text{var}(\Delta N(t)) + \lambda e^{\frac{1}{2}} \sigma_e^2 \quad \text{(A.14)}
\]

Let $\hat{\sigma}_r^2$ be the sample variance of returns and $\hat{\sigma}_\Delta N(t)$ the sample variance of the count. Plugging in $\hat{\mu}$ and $\hat{\lambda}$ from Equations (A.9) and (A.10) into Equation (A.14), we can estimate $\sigma_e^2$ by

\[
\hat{\sigma}_e^2 = \frac{\hat{\sigma}_r^2 - \hat{\mu}^2 \hat{\sigma}_\Delta N(t)}{\hat{\lambda} e^{\frac{1}{2}}} . \quad \text{(A.15)}
\]

In the following theorem, we prove that our proposed model estimators for $d = 0$ are consistent.

**Theorem A.2.** Under the null hypothesis $d = 0$, $\hat{\mu}$, $\hat{\lambda}$, $\hat{\sigma}_e$, and $\hat{c}$ are consistent estimators.

**B Simulation Methodology and Performance of Parameter Estimates**

In this section, we describe a method of simulating realizations of the log $P(t)$ process. We start with simulation of the counting process.
We approximate the intensity \( \Lambda(k) := \int_0^k \lambda e^{Z_H(s)} \, ds \) with the Riemann sum

\[
\tilde{\Lambda}(k) = \frac{1}{M} \sum_{i=1}^{kM} \lambda e^{Z_H(i/M)},
\]  

(B.1)

where \( i \) denotes the partition index, \( i = 1, 2, \ldots, kM, \ k = 1, 2, \ldots, T, \) and \( M \) is the number of partitions. Define a stationary Gaussian time series \( \{X_j\}_{j=-\infty}^{\infty} \) with mean zero and lag-\( r \) autocovariance function

\[
\gamma_x(r) = \frac{1}{2} \left[ \left| \frac{c r}{M} + 1 \right|^{2H} - 2 \left| \frac{c r}{M} \right|^{2H} + \left| \frac{c r}{M} - 1 \right|^{2H} \right].
\]  

(B.2)

Note that

\[
\left\{ \tilde{\Lambda}(k) \right\}_{k=1}^{T} \overset{d}{=} \frac{1}{M} \sum_{j=1}^{kM} \lambda e^{X_j}, \ k = 1, 2, \ldots, T.
\]  

(B.3)

We apply Davis and Harte’s (1987) method to simulate realizations of \( \{X_j\} \) and then generate \( \{\tilde{\Lambda}(k)\}_{k=1}^{T} \) by (B.3).

Given \( \tilde{\Lambda} \) (when \( d > 0 \) or when \( d = 0 \)), we simulate a sequence of independent exponential random variables \( \{\tau_j\} \) with mean \( 1/\lambda \). For each positive integer \( k \), the simulated value of the counting process \( N(k) \) is the greatest integer \( j \) such that

\[
\sum_{j=1}^{N(k)} \tau_j \leq \tilde{\Lambda}(k).
\]  

(B.4)

\( ^4 \) The counts can then be generated by

\[
\Delta N(k) = N(k) - N(k-1), \ k \geq 1.
\]  

(B.5)

We consider the special case of \( c = 1 \) for the simulations. First we simulate \( 2^p \) observations of \( \{X_j\} \), where \( p \) is a positive integer, and 40,000,000 exponential durations with \( \lambda = 128.2085 \times 10^{-6} \) to obtain the counts \( \{N(k)\} \). We then simulate a sequence of iid efficient shocks \( \{e_i\} \) from

\( ^5 \) Conditional on \( \tilde{\Lambda}(k), N(k) \) is a Poisson process with mean measure \( \tilde{\Lambda}(k) \), which represents the expected number of transactions within the interval \( \tilde{\Lambda}(k) \). To simulate the Poisson process, we first simulate iid exponential durations \( \tau_j \) with mean \( 1/\lambda \). The cumulative sum of \( \tau_j \) represents the arrival time of transactions. Hence \( N(k) \) is obtained by finding the \( k^{th} \) transaction whose arrival time is closest to \( \tilde{\Lambda}(k) \).

The model parameters are calibrated based on the mean and standard deviation of the 5-minute Boeing returns used in Cao et al. (2017), 3.846 \times 10^{-6} and 0.0012. Assume the average number of transactions per 5 minutes is 2.71.
a normal distribution with mean zero and standard deviation $\sigma_e = 0.0007289$. We assume $\mu = 1.419188 \times 10^{-6}$ and obtain the log calendar-time price process by the following equation

$$\log P(k) = \mu N(k) + \sum_{i=1}^{k} e_i,$$

which is used to generate the daily returns $r_t = \log P(t) - \log P(t-1)$, for $t = 1, 2, \ldots, T$. We aggregate $\{r_t\}$ to obtain the monthly returns $R_i = \sum_{k=(i-1)m+1}^{im} r_k$, for $i = 1, 2, \ldots, \tilde{T}$. Here we assume $m = 20$ trading days per month. We then aggregate the monthly returns $\{R_i\}$ over $\tilde{H}$ months to obtain the forward-aggregated returns $R_{i, \tilde{i}+\tilde{H}} = \sum_{j=i}^{i+\tilde{H}} R_j$. Similarly, we aggregate the squared daily returns to obtain the monthly realized variance $RV_i = \sum_{k=(i-1)m+1}^{im} r_t^2$ and the backward-aggregated realized variance $RV_{t-\tilde{H}, i} = \sum_{j=i-\tilde{H}}^{i} RV_j$. 

Figure 9 shows the estimated $\hat{\rho}_{\tilde{H}, \tilde{T}}$ for $\tilde{H} = 1, 2, \ldots, 120$ from two realizations of $\{r_t\}$ of our model. The upward trend in $\hat{\rho}_{\tilde{H}, \tilde{T}}$ as $\tilde{H}$ increases is similar to the one shown in Figure 1 of Sizova (2013), where she measured return predictability by estimating (1) the coefficient of determination $R^2$ from regressing the forward-aggregated returns on the backward-aggregated realized variance and (2) the sample correlations between the forward-aggregated returns and the backward-aggregated dividend yields using the empirical 1952-2011 AMEX/NYSE data. The similar pattern shown in our simulations and Sizova’s (2013) Figure 1 indicates that the return process generated from our model matches the return-predictability pattern observed in empirical financial data.

We estimate the parameters of our model using the simulated data and evaluate their biases with respect to the model parameters. We present our estimation results in Table 13 and Figure 10. These results indicate a larger estimation bias for a larger memory parameter $d$.

---

6Davis and Harte (1987) use the fast Fourier transform (FFT) algorithm for simulating $\{X_j\}$. To speed up the FFT computation, we simulate $2^p$ observations of $\{X_j\}$ and use $M = 50$ steps to approximate $\Lambda(k)$. This will yield approximately $T = 2^p/M$ daily returns. By assuming 20 trading days per month, we expect to have $\tilde{T} = 2^p/(20M)$ monthly returns. As a result, $\tilde{T}$ may not be an integer. For example, when $p = 22$, we obtain 4194 monthly returns. In addition, since we simulate exponential durations to obtain the counting process, the number of simulated daily returns is random, but the variation is very small. For example, we simulated $2^{19}$, $2^{20}$, and $2^{21}$ observations of $\{X_j\}$; in each case, the number of resulting daily returns was identical.
Figure 9: Return predictability $\hat{\rho}_{\tilde{H}, \tilde{T}}$ from two realizations of $\{r_t\}$.

Table 13: Estimated model parameters for $T = 83, 886$, with $d = 0, 0.15, 0.25, 0.35$, $\mu = 0.000001419188$, $\lambda = 128.2085$, $c = 1$, and $\sigma_e = 0.0007289$. The reported estimates are the average of the estimated model parameters over 500 realizations. The symbol $\ast$ indicates a statistically significant difference between the average and the true model parameter at the significance level 0.05.

| $d$   | $\hat{\mu}$     | $\hat{\lambda}$ | $\hat{\sigma_e}$   | $\hat{d}$   | $\hat{c}$    |
|-------|------------------|------------------|----------------------|-------------|--------------|
| 0.0   | 1.406521e-06     | 128.1886         | 0.0007289484         | 0.00013     | 1.0011       |
| 0.15  | 1.406393e-06     | 128.0719         | 0.0007287938         | 0.1497      | 1.0027*      |
| 0.25  | 1.406578e-06     | 127.7369         | 0.0007288298         | 0.2479*     | 1.0039*      |
| 0.35  | 1.408259e-06     | 126.5984         | 0.0007288756         | 0.3411*     | 1.0741*      |
Figure 10: Boxplots of $\hat{\mu}$, $\hat{\lambda}$, $\hat{\sigma}_e$, $\hat{c}$, and $\hat{d}$ estimated from 500 realizations of the simulated counts and return processes with size $T = 83,886$ when the population long memory parameter $d = 0, 0.15, 0.25, 0.35$.

C Proofs

Proof of Lemma 2.1

Proof. The lag-$L$ autocovariance of $r_j$ is given by

$$\text{cov}(r_0, r_L) = \text{cov}(\mathbb{E}[r_0|\Lambda], \mathbb{E}[r_L|\Lambda]) = \mu^2 \text{cov}(\mathbb{E}[\Delta N(0)|\Lambda], \mathbb{E}[\Delta N(L)|\Lambda]) = \mu^2 \lambda^2 e \int_{s=-1}^{0} \int_{t=-1}^{0} \{e^{\gamma(Z(L+t-s))} - 1\} \, dt \, ds. \quad (C.1)$$

From (C.13) and Lemma C.3, as $L \to \infty$,

$$e \int_{s=-1}^{0} \int_{t=-1}^{0} \{e^{\gamma(Z(L+t-s))} - 1\} \, dt \, ds \sim e \gamma(Z(L)).$$

It follows that $\text{cov}(r_0, r_L) \sim \mu^2 \lambda^2 \gamma(Z(L))$. \qed
Proof of Corollary 2.2

Proof. The variance of $r_j$ can be represented as

\[
\text{var}(r_j) = \text{var}(E[r_j|\Lambda]) + E[\text{var}(r_j|\Lambda)]
\]

\[
= \text{var} \left( \mu \lambda \int_t^{t-1} e^{ZH(s)} ds \right) + E[\text{var}(r_j|\Lambda)]
\]

\[
= \mu^2 \lambda^2 e^1 \int_{s=-1}^0 \int_{t=-1}^0 \{ e^{-z(s-t)} - 1 \} dt ds + \lambda e^{1/2}(\mu^2 + \sigma^2_e) . \tag{C.2}
\]

\[\square\]

Lemma C.1. \( \text{cov} \left( R_{i,i-H}, RV_{i-H,i} \right) \) is given by Equation (3.4) and is positive.

Proof. By (3.1) and (3.2),

\[
R_{i,i-H} = \sum_{i=tm+1}^{(i+H)m} r_j = \sum_{j=tm+1}^{(i+H)m} \left[ \mu \Delta N(j) + \sum_{k=N(j-1)+1}^{N(j)} e_k \right], \tag{C.3}
\]

\[
RV_{i-H,i} = \sum_{j=(i-H)m+1}^{im} r_j^2 = \sum_{j=(i-H)m+1}^{im} \left[ \mu \Delta N(j) + \sum_{k=N(j-1)+1}^{N(j)} e_k \right]^2 . \tag{C.4}
\]

Hence

\[
\text{cov} \left( R_{i,i-H}, RV_{i-H,i} \right) = \text{cov} \left( \sum_{i=tm+1}^{(i+H)m} r_j , \sum_{j=tm+1}^{(i+H)m} r_j^2 \right)
\]

\[
= \mathbb{E} \left[ \sum_{j=tm+1}^{(i+H)m} \sum_{i=tm+1}^{j} r_j r_{i-Hm} \right] - \mathbb{E} \left[ \sum_{j=tm+1}^{(i+H)m} r_j \right] \mathbb{E} \left[ \sum_{j=tm+1}^{(i+H)m} r_{j-Hm}^2 \right]
\]

\[
= \sum_{j=Hm+1}^{2Hm} \sum_{i=Hm+1}^{2Hm} \text{cov}(r_j, r_{i-Hm})
\]

\[
= \sum_{L=1}^{Hm} \text{cov}(r_L, r_{Hm}^2) L + \sum_{L=Hm+1}^{2Hm-1} \left( 2Hm - L \right) \text{cov}(r_L, r_{Hm}^2) \tag{C.5}
\]

where \( L = j - (i - Hm) \).

Conditional on \( \Lambda \), \( N(t) \) is a poisson process, so its increments are independent of each other.
Since $L > 0$, $\text{cov}(\Delta N(L), (\Delta N(0))^2 | \Lambda) = 0$. The efficient shock $e_k$ is independent of the counting process. These properties imply

$$\mathbb{E} \left[ \text{cov}(\Delta N(L), \left( \sum_{k=N(-1)+1}^{N(0)} e_k \right)^2 | \Lambda) \right] = 0$$

and

$$\mathbb{E}[\text{cov}(r_L, r_0^2 | \Lambda) = 0] .$$

Hence

$$\text{cov} (r_L, r_0^2) = \text{cov} \left( \mathbb{E}[r_L | \Lambda], \mathbb{E}[r_0^2 | \Lambda] \right)$$

$$= \mu^2 \text{cov} \left( \mathbb{E}[\Delta N(L) | \Lambda], \mathbb{E}[\text{cov}(\Delta N(0))^2 | \Lambda) \right)$$

$$+ \mu \text{cov} \left( \mathbb{E}[\Delta N(L) | \Lambda], \mathbb{E} \left[ \left( \sum_{k=N(-1)+1}^{N(0)} e_k \right)^2 | \Lambda) \right] \right)$$

(C.6)

First, we consider

$$\text{cov} \left( \mathbb{E}[\Delta N(L) | \Lambda], \mathbb{E}[\text{cov}(\Delta N(0))^2 | \Lambda) \right)$$

$$= \text{cov} \left( \lambda \int_{t=L-1}^{L} e^{Z_H(t)} dt, \lambda \int_{t=L-1}^{0} e^{Z_H(t)} dt + \lambda^2 \int_{t=L-1}^{0} \int_{s=L-1}^{0} e^{Z_H(s)} e^{Z_H(t)} ds dt \right)$$

$$= \lambda^2 \text{cov} \left( \int_{s=L-1}^{L} e^{Z_H(s)} ds, \int_{t=L-1}^{0} e^{Z_H(t)} dt \right)$$

$$+ \lambda^3 \text{cov} \left( \int_{s=L-1}^{L} e^{Z_H(s)} ds, \int_{t=L-1}^{0} \int_{s=L-1}^{0} e^{Z_H(s)} e^{Z_H(t)} ds dt \right)$$

(C.7)

For all non-negative integer $L$, define

$$A_L = \int_{s=L-1}^{0} e^{Z_H(s)} ds = \int_{s=L-1}^{L} e^{Z_H(s)} ds$$

(C.8)

$$B_L = \int_{s=L-1}^{0} \int_{t=L-1}^{0} e^{Z_H(s)} e^{Z_H(t)} ds dt = \int_{s=L-1}^{L} \int_{t=L-1}^{L} e^{Z_H(s)} e^{Z_H(t)} ds dt$$

(C.9)

Hence Equation (C.7) can also be represented as

$$\text{cov} \left( \mathbb{E}[\Delta N(L) | \Lambda], \mathbb{E}[\text{cov}(\Delta N(0))^2 | \Lambda) \right) = \lambda^2 \text{cov} (A_L, A_0) + \lambda^3 \text{cov} (A_L, B_0) .$$

(C.10)

We will use the following property to evaluate (C.10). Let $(X, Y)$ be a Gaussian vector such that $\mathbb{E}[X] = \mathbb{E}[Y] = 0$, $\text{var}(X) = \sigma_X^2$, $\text{var}(Y) = \sigma_Y^2$ and $\text{cov}(X, Y) = \rho$. Note that

$$\text{cov}(e^X, e^Y) = e^{\sigma_X^2/2 + \sigma_Y^2/2}(e^\rho - 1) .$$

(C.11)
This yields
\[
\text{cov} \left( e^{Z_H(s)}, e^{Z_H(t)}e^{Z_H(u)} \right) = e^{1/2+(1/2+1/2+\gamma_Z(t-u))} \left( e^{\gamma_Z(s-t)+\gamma_Z(s-u)} - 1 \right) \\
= e^{3/2} e^{\gamma_Z(t-u)} \left[ e^{\gamma_Z(s-t)+\gamma_Z(s-u)} - 1 \right] \quad \text{(C.12)}
\]

By Fubini’s Theorem as well as (C.11) and (C.12), Equation (C.7) becomes
\[
\text{cov} \left( \mathbb{E}[\Delta N(L)|A], \mathbb{E}[\Delta N(0)^2|A] \right) = \lambda^2 \int_{s=L-1}^{0} \int_{t=0}^{0} \text{cov}(e^{Z_H(s)}, e^{Z_H(t)})dt ds \\
+ \lambda^3 e^{3/2} \int_{s=L-1}^{0} \int_{t=0}^{0} \int_{u=0}^{0} e^{\gamma_Z(t-u)} \left[ e^{\gamma_Z(s-t)+\gamma_Z(s-u)} - 1 \right] du dt ds \\
= \lambda^2 \int_{s=L-1}^{0} \int_{t=0}^{0} \left\{ e^{\gamma_Z(L+s-t)} - 1 \right\} dt ds \\
+ \lambda^3 e^{3/2} \int_{s=L-1}^{0} \int_{t=0}^{0} \int_{u=0}^{0} e^{\gamma_Z(t-u)} \left[ e^{\gamma_Z(s-t)+\gamma_Z(s-u+L)} - 1 \right] du dt ds.
\]

It follows that
\[
\text{cov}(A_L, A_0) = e \int_{s=1}^{0} \int_{t=0}^{0} \left\{ e^{\gamma_Z(L+s-t)} - 1 \right\} dt ds \quad \text{(C.13)}
\]

\[
\text{cov}(A_L, B_0) = e^{3/2} \int_{s=1}^{0} \int_{t=0}^{0} \int_{u=0}^{0} e^{\gamma_Z(t-u)} \left[ e^{\gamma_Z(s-t)+\gamma_Z(s-u+L)} - 1 \right] du dt ds \\
= \text{cov}(A_0, B_L). \quad \text{(C.14)}
\]

Next, we consider
\[
\text{cov} \left( \mathbb{E}[\Delta N(L)|A], \mathbb{E} \left[ \left( \sum_{k=N(1)+1}^{N(0)} e_k \right)^2 |A \right] \right) = \lambda^2 \text{cov} \left( \int_{s=L-1}^{0} e^{Z_H(s)} ds, \sigma^2 \int_{s=1}^{0} e^{Z_H(s)} ds \right) \\
= \lambda^2 \sigma^2 e^{3/2} \int_{s=L-1}^{0} \int_{t=1}^{0} \text{cov}(e^{Z_H(s)}, e^{Z_H(t)}) ds dt \\
= \sigma^2 \lambda^2 e^{3/2} \int_{s=L-1}^{0} \int_{t=1}^{0} \left\{ e^{\gamma_Z(L+s-t)} - 1 \right\} dt ds \\
= \sigma^2 \lambda^2 \text{cov}(A_L, A_0). \quad \text{(C.15)}
\]

Combining Equations (C.10) and (C.15), \( \forall L > 0 \), we obtain
\[
\text{cov}(r_L, r_0^2) = \mu^3 \left( \lambda^3 \text{cov}(A_0, B_L) + \lambda^2 \text{cov}(A_L, A_0) \right) + \mu \sigma^2 \lambda^2 \text{cov}(A_L, A_0). \quad \text{(C.16)}
\]

Since \( \gamma_Z(r) > 0 \forall r \in \mathbb{R} \) (see Lemma C.6), it follows that \( \text{cov}(A_0, B_L) > 0 \) and \( \text{cov}(A_L, A_0) > 0 \) and thus \( \text{cov}(r_L, r_0^2) > 0 \). From Equation (3.4), it follows that \( \text{cov}(R_{t+\tilde{H}, L}, RV_{t-H,t}) > 0 \).
Corollary C.2. For model in Equation (2.10), \( \text{cov} \left( R_{i,i+H}, RV_{i-H,i} \right) > 0 \).

Proof. The return process is then given by

\[
r_t = \mu_1 \Delta N_1(t) + \mu_2 \Delta N_2(t) + \sum_{k=N_1(i-1)+1}^{N_1(t)} e_{1,k} + \sum_{k=N_2(i-1)+1}^{N_2(t)} e_{2,k}
\]  

(C.17)

Under the new model, Equations (C.3) and (C.4) become

\[
R_{i,i+1} = \sum_{j=bm+1}^{(i+1)m} \left[ \mu_1 \Delta N_1(t) + \mu_2 \Delta N_2(t) + \sum_{k=N_1(i-1)+1}^{N_1(t)} e_{1,k} + \sum_{k=N_2(i-1)+1}^{N_2(t)} e_{2,k} \right]
\]  

(C.18)

\[
RV_{i-1,i} = \sum_{j=(i-1)m+1}^{im} \left[ \mu_1 \Delta N_1(t) + \mu_2 \Delta N_2(t) + \sum_{k=N_1(i-1)+1}^{N_1(t)} e_{1,k} + \sum_{k=N_2(i-1)+1}^{N_2(t)} e_{2,k} \right]^2
\]  

(C.19)

Using the expression for the \( r_i \) in Equation (C.17) and the assumptions we make for the efficient shocks and the counting processes, we obtain

\[
\text{cov} \left( r_0, r_L^2 \right) = \mu_1^3 \text{cov} \left( \Delta N_1(0), (\Delta N_1(L))^2 \right)
\]  

\[
+ \mu_1 \text{cov} \left( \Delta N_1(0), \left( \sum_{k=N_1(L-1)+1}^{N_1(L)} e_{1,k} \right)^2 \right)
\]  

\[
+ \mu_2^3 \text{cov} \left( \Delta N_2(0), (\Delta N_2(L))^2 \right)
\]  

\[
+ \mu_2 \text{cov} \left( \Delta N_2(0), \left( \sum_{k=N_2(L-1)+1}^{N_2(L)} e_{2,k} \right)^2 \right)
\]  

\[
+ 2\mu_1 \mu_2 \text{cov} \left( \Delta N_1(0), \Delta N_1(L) \Delta N_2(L) \right)
\]  

\[
+ 2\mu_2 \mu_1^2 \text{cov} \left( \Delta N_2(0), \Delta N_1(L) \Delta N_2(L) \right)
\]  

\[
+ 2\mu_1 \text{cov} \left( \Delta N_1(0), \left( \sum_{k=N_1(L-1)+1}^{N_1(L)} e_{1,k} \right) \times \left( \sum_{k=N_2(L-1)+1}^{N_2(L)} e_{2,k} \right) \right)
\]  

\[
+ 2\mu_2 \text{cov} \left( \Delta N_2(0), \left( \sum_{k=N_1(L-1)+1}^{N_1(L)} e_{1,k} \right) \times \left( \sum_{k=N_2(L-1)+1}^{N_2(L)} e_{2,k} \right) \right)
\]  

(C.20)
Note

\[
\text{cov} \left( \Delta N_1(0), \left( \sum_{k=N_1(L-1)+1}^{N_1(L)} e_{1,k} \right) \times \left( \sum_{k=N_2(L-1)+1}^{N_2(L)} e_{2,k} \right) \right)
\]

\[
= \text{cov} \left( \mathbb{E}[\Delta N_1(0) | \Lambda], \mathbb{E} \left[ \sum_{k=N_1(L-1)+1}^{N_1(L)} \sum_{l=N_2(L-1)+1}^{N_2(L)} e_{1,k} e_{2,l} | \Lambda \right] \right)
\]

\[
+ \mathbb{E} \left[ \text{cov} (\Delta N_1(0) | \Lambda), \text{cov} \left( \sum_{k=N_1(L-1)+1}^{N_1(L)} \sum_{l=N_2(L-1)+1}^{N_2(L)} e_{1,k} e_{2,l} | \Lambda \right) \right]
\]

\[
= 0.
\]

The same conclusion for

\[
\text{cov} \left( \Delta N_2(0), \left( \sum_{k=N_1(L-1)+1}^{N_1(L)} e_{1,k} \right) \times \left( \sum_{k=N_2(L-1)+1}^{N_2(L)} e_{2,k} \right) \right) = 0.
\]

\[
\text{cov} (\Delta N_1(0), \Delta N_1(L) \Delta N_2(L)) = \text{cov} (\mathbb{E}[\Delta N_1(0) | \Lambda], \mathbb{E}[\Delta N_1(L) \Delta N_2(L) | \Lambda])
\]

\[
+ \mathbb{E} \left[ \text{cov}(\Delta N_1(0), \Delta N_1(L) \Delta N_2(L)) | \Lambda \right]
\]

\[
= \lambda^3 \text{cov} \left( \int_{s=0}^{0} e^{Z_H(s)} ds, \int_{t=L-1}^{L} \int_{s=L-1}^{L} e^{Z_H(s)} e^{Z_H(t)} ds dt \right)
\]

\[
= \lambda^3 \int_{s=-1}^{0} \int_{t=L-1}^{L} \int_{u=L-1}^{L} \text{cov} (e^{Z_H(s)}, e^{Z_H(t)} e^{Z_H(u)}) du dt ds
\]

\[
= \text{cov}(A_0, B_L)
\]

Same for \(\text{cov} (\Delta N_2(0), \Delta N_1(L) \Delta N_2(L)) = \text{cov}(A_0, B_L)\).

Therefore, Equation \((\text{C.20})\) becomes

\[
\text{cov} (r_0, r_L^2) = (\mu_1^3 + \mu_2^3) (\lambda^3 \text{cov}(A_0, B_L) + \lambda^2 \text{cov}(A_0, A_L)) + 2 \mu_1 \mu_2 (\mu_1 + \mu_2) \lambda^3 \text{cov}(A_0, B_L)
\]

\[
+ 2(\mu_1 + \mu_2) \lambda^2 \sigma^2 \text{cov}(A_L, A_0)
\]

\[
= (\mu_1 + \mu_2)(\mu_1^2 + \mu_2^2 + \mu_1 \mu_2) \lambda^3 \text{cov}(A_0, B_L) + (\mu_1^3 + \mu_2^3) \lambda^2 \text{cov}(A_0, A_L)
\]

\[
+ 2(\mu_1 + \mu_2) \lambda^2 \sigma^2 \text{cov}(A_L, A_0).
\]

Note \((\mu_1 + \mu_2)(\mu_1^2 + \mu_2^2 + \mu_1 \mu_2) = (\mu_1 + \mu_2)^3 - \mu_1 \mu_2 (\mu_1 + \mu_2) > 0, \) \(\mu_1^3 + \mu_2^3 = (\mu_1 + \mu_2)(\mu_1^2 + \mu_2^2 - \mu_1 \mu_2) > 0. \) This proves \(\text{cov} (r_0, r_L^2) > 0. \) \(\square\)
Lemma C.3. Let $A_L$ and $B_L$ be defined in (C.8) and (C.9) respectively. As $L \to \infty$, 
$\text{cov}(A_L, A_0) \sim e \gamma_L(L)$, $\text{cov}(A_0, B_L) \sim 2e^3 \gamma_L(L) \int_{s=1}^{0} \int_{t=-1}^{0} e^{\gamma_L(s-u)} du ds$, and
\[
\text{cov}(B_0, B_L) \sim 4e^2 \gamma_L(L) \int_{s=-1}^{0} \int_{t=-1}^{0} \int_{u=-1}^{0} e^{\gamma_L(s-t)} e^{\gamma_L(u-v)} dv du ds. \quad (C.21)
\]

Proof. It can be shown that
\[
\text{cov}(A_0, A_L) = e \int_{s=-1}^{0} \int_{t=-1}^{0} \text{cov}(e^{Z_{H(s+L)}}, e^{Z_{H(t)}}) dt ds
\]
\[
= e \int_{s=-1}^{0} \int_{t=-1}^{0} \{e^{\gamma_L(L+t-s)} - 1\} dt ds
\]
\[
= e \gamma_L(L) \int_{u=-1}^{1} (1 - |u|) \left\{ \frac{e^{\gamma_L(L+u)} - 1}{\gamma_L(L)} \right\} du \quad (C.22)
\]

The ratio $\frac{e^{\gamma_L(L+u)} - 1}{\gamma_L(L)}$ is uniformly bounded for all $u \in (-1, 1)$ and by condition (2.3), the ratio converges to one as $L$ goes to $\infty$. By the dominated convergence theorem, the integral converges to $\int_{-1}^{1} (1 - |t|) dt = 1$. Therefore, as $L \to \infty$,
\[
\text{cov}(A_0, A_L) \sim e \gamma_L(L). \quad (C.23)
\]

Next,
\[
\text{cov}(A_L, B_0) = e^3 \int_{t=-1}^{0} \int_{s=-1}^{0} \int_{u=-1}^{0} \text{cov}(e^{Z_{H(t+L)}}, e^{Z_{H(s)}e^{Z_{H(u)}}}) du ds dt
\]
\[
= e^3 \int_{t=-1}^{0} \int_{s=-1}^{0} \int_{u=-1}^{0} e^{\gamma_L(s-u)} \left[ e^{\gamma_L(L+t-s)} e^{\gamma_L(L+t-u)} - 1 \right] du ds dt
\]
\[
= e^3 \gamma_L(L) \int_{t=-1}^{0} \int_{s=-1}^{0} \int_{u=-1}^{0} e^{\gamma_L(s-u)} \left\{ \frac{e^{\gamma_L(L+t-s)} e^{\gamma_L(L+t-u)} - 1}{\gamma_L(L)} \right\} du ds dt
\]

Again condition (2.3) implies that the ratio $\frac{e^{\gamma_L(L+t-s)} e^{\gamma_L(L+t-u)} - 1}{\gamma_L(L)}$ is uniformly bounded for $-1 \leq u \leq 0$, $-1 \leq v \leq 0$, and $-1 \leq t \leq 0$ and converges to 2 as $L$ goes to $\infty$. The function $e^{\gamma_L(s-u)} \left\{ \frac{e^{\gamma_L(L+t-s)} e^{\gamma_L(L+t-u)} - 1}{\gamma_L(L)} \right\}$ is positive and is integrable with respect to the Lebesgue product measure. By the Fubini’s theorem and the dominated convergence theorem, as $L \to \infty$,
\[
\text{cov}(A_L, B_0) \sim 2e^3 \gamma_L(L) \int_{s=-1}^{0} \int_{u=-1}^{0} e^{\gamma_L(s-u)} du ds \quad (C.24)
\]
Finally, we can express
\[ \text{cov}(B_L, B_0) = \int_{s=-1}^{0} \int_{t=-1}^{0} \int_{u=-1}^{0} \int_{v=-1}^{0} \text{cov} \left( e^{Z_H(s+L)} e^{Z_H(t+L)}, e^{Z_H(v)} e^{Z_H(u)} \right) dudvds \]
\[ = e^2 \int_{s=-1}^{0} \int_{t=-1}^{0} \int_{u=-1}^{0} \int_{v=-1}^{0} e^{\gamma z(s-t)} e^{\gamma z(u-v)} \left[ e^{\gamma z(L+s-v)} e^{\gamma z(L+s-u)} e^{\gamma z(L+t-v)} e^{\gamma z(L+t-u)} - 1 \right] dudvds \] (C.25)

Applying a similar argument as for \( \text{cov}(A_L, B_0) \) and the dominated convergence theorem, we can show that as \( L \to \infty \),

\[ \text{cov}(B_L, B_0) \sim 4e^2 \gamma z(L) \int_{s=-1}^{0} \int_{t=-1}^{0} \int_{u=-1}^{0} \int_{v=-1}^{0} e^{\gamma z(s-t)} e^{\gamma z(u-v)} dudvds . \] (C.26)

We will use Lemma [C.1] and Lemma [C.3] for the proof of Theorem 2.3.

**Proof of Theorem 2.3**

*Proof.* For \( L > 0 \), conditional on \( \Lambda \), \( \Delta N(0) \) and \( \Delta N(L) \) are independent. Therefore,

\[ \text{cov}(r_0^2, r_L^2) = \text{cov} \left( E[r_0^2|\Lambda], E[r_L^2|\Lambda] \right) \]
\[ = \mu^4 \text{cov} \left( E[(\Delta N(0))^2|\Lambda], E[(\Delta N(L))^2|\Lambda] \right) + \mu^2 \sigma^2_\epsilon \text{cov} \left( E[\Delta N(0)|\Lambda], E[(\Delta N(L))^2|\Lambda] \right) \]
\[ + \mu^2 \sigma^2_\epsilon \text{cov} \left( E[\Delta N(L)|\Lambda], E[(\Delta N(0))^2|\Lambda] \right) + \sigma^4_\epsilon \text{cov} \left( E[\Delta N(0)|\Lambda], E[\Delta N(L)|\Lambda] \right) \] (C.28)

where

\[ \text{cov} \left( E[(\Delta N(0))^2|\Lambda], E[(\Delta N(L))^2|\Lambda] \right) = \lambda^3 \text{cov}(A_0, A_L) + \lambda^3 \text{cov}(A_L, B_0) + \lambda^3 \text{cov}(A_0, B_L) + \lambda^4 \text{cov}(B_L, B_0) \]
\[ = \lambda^3 \text{cov}(A_0, A_L) + 2\lambda^3 \text{cov}(A_L, B_0) + \lambda^4 \text{cov}(B_L, B_0) \] (C.29)

\[ \text{cov} \left( E[\Delta N(0)|\Lambda], E[\Delta N(L)^2|\Lambda] \right) \]
\[ = \text{cov} \left( E[\Delta N(L)|\Lambda], E[\Delta N(0)^2|\Lambda] \right) , \] (C.30)
and
\[
\text{cov} \left( \mathbb{E}[\Delta N(0)|\Lambda] , \mathbb{E}[\Delta N(L)|\Lambda] \right) = \lambda^2 \text{cov}(A_0, A_L) .
\] (C.31)

Using (C.10), (C.29), (C.30), and (C.31), \( \text{cov}(r_0^2, r_L^2) \) can be expressed as
\[
\text{cov}(r_0^2, r_L^2) = (\mu^2 + \sigma_e^2)^2 \lambda^2 \text{cov}(A_0, A_L) + 2\mu^2 \lambda^3 (\mu^2 + \sigma_e^2) \text{cov}(A_0, B_L) + \mu^4 \lambda^4 \text{cov}(B_0, B_L) .
\] (C.32)

Applying Lemma C.3 to (C.32), it follows that as \( L \to \infty \),
\[
\lim_{L \to \infty} \tilde{\gamma}(L) = C .
\]

where
\[
C = \mu^4 C_1 + 2\mu^2 \sigma^2 C_2 + \sigma^4 \lambda^2 e ,
\] \hspace{1cm} (C.33)
\[
C_1 = \lambda^2 e \left( 1 + 4\lambda e^{\frac{d}{2}} \int_{s=-1}^{0} \int_{u=-1}^{0} e^{\gamma_Z(s-u)} duds \right)
+ 4\lambda^2 e \int_{s=-1}^{0} \int_{t=-1}^{0} \int_{u=-1}^{0} \int_{v=-1}^{0} e^{\gamma_Z(s-t)} e^{\gamma_Z(u-v)} dudvtds ,
\] \hspace{1cm} (C.34)
\[
C_2 = \lambda^2 e \left( 2\lambda e^{\frac{d}{2}} \int_{s=-1}^{0} \int_{u=-1}^{0} e^{\gamma_Z(s-u)} duds + 1 \right) .
\] \hspace{1cm} (C.35)

Lemma C.4.
\[
\operatorname{var} \left( R_V_{l-H,m} \right) \sim \tilde{C}_{\mu,\lambda,d,\sigma_e} \frac{\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{d(2d + 1)}
\]
where \( \tilde{C}_{\mu,\lambda,d,\sigma_e} > 0 \) is a function of \( \mu, \lambda, d, \) and \( \sigma_e \) defined in (C.36).

Proof. Let \( \tilde{\gamma}(k) \) be the lag-k autocovariance function of squared return. In Theorem 2.3 we prove that as \( k \to \infty \)
\[
\text{cov}(r_0^2, r_k^2) \sim (\mu^4 C_1 + 2\mu^2 \sigma^2 C_2 + \sigma^4 \lambda^2 e) \gamma_Z(k) .
\]
Therefore, \( \tilde{\gamma}(k) \) can be represented as
\[
\tilde{\gamma}(k) = \tilde{C}_{\mu,\lambda,d,\sigma_e} \ell(k) k^{2d-1} ,
\]
44
where \( \ell(k) \) is a slowly varying function of fractional Gaussian which satisfies \( \lim_{k \to \infty} \ell(k) = \frac{2d+1}{4d} \) and

\[
\tilde{C}_{\mu,\lambda,d,\sigma_e} = \mu^4 C_1 + 2\mu^2 \sigma^2 C_2 + \sigma^4 \lambda^2 e. \tag{C.36}
\]

Note that

\[
\text{var} \left( RV_{\tilde{t}-\tilde{H},\tilde{t}} \right) = \text{cov} \left( \sum_{j=(\tilde{t}-\tilde{H})m+1}^{\tilde{t}m} r_j^2, \sum_{j=(\tilde{t}-\tilde{H})m+1}^{\tilde{t}m} r_j^2 \right)
\]

\[
= \sum_{j=(\tilde{t}-\tilde{H})m+1}^{\tilde{t}m} \sum_{i=(\tilde{t}-\tilde{H})m+1}^{\tilde{t}m} \text{cov} \left( r_j^2, r_i^2 \right)
\]

\[
= \tilde{H}m-1 \sum_{k=-(\tilde{H}m-1)}^{\tilde{H}m-1} \left( \tilde{H}m - |k| \right) \tilde{\gamma}(k)
\]

\[
= \tilde{H}m-1 \sum_{k=-(\tilde{H}m-1)}^{\tilde{H}m-1} (\tilde{H}m)\tilde{\gamma}(k) - \sum_{k=-(\tilde{H}m-1)}^{\tilde{H}m-1} |k|\tilde{\gamma}(k)
\]

\[
= (\tilde{H}m)\text{var}(r_0^2) + 2(\tilde{H}m) \sum_{k=1}^{\tilde{H}m-1} \tilde{\gamma}(k) - 2 \sum_{k=1}^{\tilde{H}m-1} k\tilde{\gamma}(k) \tag{C.37}
\]

Since \( 0 < d < \frac{1}{2} \), \( (\tilde{H}m)\text{var}(r_0^2) = o(\tilde{H}^{2d+1}) \). Applying Proposition 2.2.1 again from Pipiras and Taqqu (2017) for \( \sum_{k=1}^{n} \tilde{\gamma}(k) \) as \( n \to \infty \), \( \sum_{k=1}^{n} \tilde{\gamma}(k) \sim \frac{\ell(n)n^{2d}}{2d} \). We obtain

\[
\text{var} \left( RV_{\tilde{t}-\tilde{H},\tilde{t}} \right) \sim \tilde{C}_{\mu,\lambda,d,\sigma_e} \left\{ \frac{2\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{2d} - \frac{2\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{2d+1} \right\}
\]

\[
= \tilde{C}_{\mu,\lambda,d,\sigma_e} \frac{\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{d(2d+1)} \tag{C.38}
\]

Lemma C.5.

\[
\text{var} \left( R_{\tilde{t},\tilde{t}+\tilde{H}} \right) \sim C_{\mu,\lambda} \frac{\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{d(2d+1)},
\]

where \( C_{\mu,\lambda} = \mu^2 \lambda^2 e \).

Proof. Let \( \gamma_r(k) = \text{cov}(r_0, r_k) \) be the lag-k autocovariance function of returns. By Equation (2.12) and the convergence result in (C.22), as \( k \to \infty \)

\[
\text{cov}(r_0, r_k) \sim \mu^2 \lambda^2 e \gamma_z(k).
\]
Therefore, \( \gamma_r(k) \) has a representation of a slowly varying function as

\[
\gamma_r(k) = C_{\mu,\lambda} \ell(k) k^{2d - 1},
\]

where \( C_{\mu,\lambda} = \mu^2 \lambda^2 e \) and \( \ell(k) \) is a slowly varying function of fractional Gaussian which satisfies \( \lim_{k \to \infty} \ell(k) = \frac{2d+1}{4d} \).

Note

\[
\text{var} \left( R_{\tilde{t}, \tilde{t}+\tilde{H}} \right) = \text{cov} \left( R_{\tilde{t}, \tilde{t}+\tilde{H}} , R_{\tilde{t}, \tilde{t}+\tilde{H}} \right) = \text{cov} \left( \sum_{j=\tilde{t}m+1}^{\tilde{H}m-1} r_j , \sum_{i=\tilde{t}m+1}^{\tilde{H}m-1} r_i \right) = (\tilde{H}m) \text{var}(r_0) + 2 \sum_{k=1}^{\tilde{H}m-1} (\tilde{H}m - k) \gamma_r(k) = (\tilde{H}m) \text{var}(r_0) + 2 \sum_{k=1}^{\tilde{H}m-1} (\tilde{H}m) \gamma_r(k) - 2 \sum_{k=1}^{\tilde{H}m-1} k \gamma_r(k)
\]

Since \( 0 < d < \frac{1}{2} \), \( (\tilde{H}m) \text{var}(r_0) = o(\tilde{H}^{2d+1}) \). Applying Proposition 2.2.1 again from Pipiras and Taqqu (2017) for \( \sum_{k=1}^{n} \gamma_r(k) \) when \( n \to \infty \), we obtain

\[
\text{var} \left( R_{\tilde{t}, \tilde{t}+\tilde{H}} \right) \sim C_{\mu,\lambda} \left\{ \frac{2\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{2d} - \frac{2\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{2d+1} \right\} = C_{\mu,\lambda} \frac{\ell(\tilde{H}m)(\tilde{H}m)^{2d+1}}{d(2d+1)}. \tag{C.40}
\]

We will use Lemmas C.4 and C.5 for the proof of Theorem 3.1.

Proof of Theorem 3.1

Proof. Equation (C.16) and the convergence results of (C.23) and (C.24) imply that as \( k \to \infty \),

\[
\text{cov}(r_k, r_0^2) \sim C_{\mu,\lambda,d,\sigma_e} \gamma_Z(k)
\]
where

\[ C_{\mu,\lambda,d,\sigma_e} = \mu^3 C_2 + \mu \sigma_e^2 \lambda^2 e \]  

(C.41)

Define \( \gamma_c(k) = \text{cov}(r_k, r_0^2) \). It can be represented as

\[ \gamma_c(k) = C_{\mu,\lambda,d,\sigma_e} \ell(k) k^{2d-1}, \]

where \( \ell(k) \) is a slowly varying function of fractional Gaussian noise which satisfies \( \lim_{k \to \infty} \ell(k) = \frac{2d+1}{4d} \).

From Equation (3.4), \( \text{cov}(R_{t,t+\tilde{H}}, RV_{t-\tilde{H},t}) \) can be expressed as

\[
\text{cov}(R_{t,t+\tilde{H}}, RV_{t-\tilde{H},t}) = \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) + \sum_{k=\tilde{H} m+1}^{2 \tilde{H} m-1} \left(2 \tilde{H} m - k\right) \gamma_c(k) \\
= \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) + \sum_{k=1}^{2 \tilde{H} m-1} \left(2 \tilde{H} m - k\right) \gamma_c(k) - \sum_{k=1}^{\tilde{H} m} (2 \tilde{H} m - k) \gamma_c(k) \\
= \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) + 2(\tilde{H} m) \sum_{k=1}^{2 \tilde{H} m-1} \gamma_c(k) - 2(\tilde{H} m) \sum_{k=1}^{\tilde{H} m} \gamma_c(k) \\
- \sum_{k=1}^{2 \tilde{H} m-1} k \gamma_c(k) + \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) \\
\approx \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) + \sum_{k=1}^{2 \tilde{H} m-1} k \gamma_c(k) \\
\approx \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) + \sum_{k=1}^{2 \tilde{H} m-1} k \gamma_c(k) \\
= \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) + \sum_{k=1}^{\tilde{H} m} k \gamma_c(k) \]

(C.42)

Applying Proposition 2.2.1 from Pipiras and Taqqu (2017) in which they showed that as \( n \to \infty \),

\[
\sum_{k=1}^{n} \gamma_c(k) \sim C_{\mu,\lambda,d,\sigma_e} \frac{\ell(n)n^{2d}}{2d},
\]

we obtain

\[
\text{cov}(R_{t,t+\tilde{H}}, RV_{t-\tilde{H},t}) \sim \left\{ \frac{\ell(\tilde{H} m)(\tilde{H} m)^{2d+1}}{2d+1} + \frac{2(\tilde{H} m)\ell(2\tilde{H} m-1)(2\tilde{H} m-1)^{2d}}{2d} \\
- \frac{(2\tilde{H} m)\ell(\tilde{H} m)(\tilde{H} m)^{2d}}{2d} \\
- \frac{\ell(2\tilde{H} m-1)(2\tilde{H} m-1)^{2d+1}}{2d+1} + \frac{\ell(\tilde{H} m)(\tilde{H} m)^{2d+1}}{2d+1} \right\} C_{\mu,\lambda,d,\sigma_e} \\
\sim C_{\mu,\lambda,d,\sigma_e} \frac{\ell(\tilde{H} m)(\tilde{H} m)^{2d+1}}{d(2d+1)} [2^{2d} - 1].
\]  

(C.43)
Combining Lemmas C.4, C.5, and Equation (C.43), we show that as $\tilde{H} \to \infty$
\[
\text{corr} \left( R_{i,i+\tilde{H}}, RV_{i-\tilde{H},i} \right) \to \frac{C_{\mu,\lambda,d,\sigma_e}}{\sqrt{\tilde{C}_{\mu,\lambda,d,\sigma_e} C_{\mu,\lambda}}} (2^{2d} - 1), \quad (C.44)
\]

Following the definitions of $C_1$, $C_2$, $\tilde{C}_{\mu,\lambda,d,\sigma_e}$, and $C_{\mu,\lambda,d,\sigma_e}$ in (C.34), (C.35), (C.36), (C.41) and Lemma C.5 it can be shown that
\[
(C_{\mu,\lambda,d,\sigma_e})^2 = \tilde{C}_{\mu,\lambda,d,\sigma_e} C_{\mu,\lambda},
\]
Therefore,
\[
\text{corr} \left( R_{i,i+\tilde{H}}, RV_{i-\tilde{H},i} \right) \to (2^{2d} - 1).
\]

Proof of Lemma A.1

Proof. For $L > 0$, conditional on $\Lambda$, $\Delta N(0)$ and $\Delta N(L)$ are independent. Therefore, the lag-L autocovariance of $\Delta N(t)$ is equal to
\[
\text{cov} \left( \mathbb{E}[\Delta N(L)|\Lambda], \mathbb{E}[(\Delta N(0)|\Lambda)] \right) = \lambda^2 \text{cov} \left( \int_{s=L-1}^{L} e^{Z_H(s)}ds, \int_{t=-1}^{0} e^{Z_H(t)}dt \right)
\]
\[
= \lambda^2 \int_{s=L-1}^{L} \int_{t=-1}^{0} \text{cov}(e^{Z_H(s)}, e^{Z_H(t)})dt ds
\]
\[
= \lambda^2 e \int_{s=-1}^{0} \int_{t=-1}^{0} \{e^{\gamma z(L+s-t)} - 1\} dt ds
\]

We can express
\[
\text{var} (\Delta N(t)) = \text{var} (\mathbb{E}[\Delta N(t)|\Lambda]) + \mathbb{E}[\text{var}(\Delta N(t)|\Lambda)]
\]
\[
= \lambda e^{\frac{1}{2}} + \lambda^2 e \int_{s=-1}^{0} \int_{t=-1}^{0} \{e^{\gamma z(s-t)} - 1\} dt ds
\]
Lemma C.6. Let $Z_H(t) = B_H(ct) - B_H(ct - 1)$, where $B_H(t)$ is the fractional Brownian motion with Hurst index $\frac{1}{2} < H < 1$, and $c > 0$. The lag-$r$ autocovariance function $\gamma_Z(r)$ of $Z_H$ is equal to

$$\gamma_Z(r) = \frac{1}{2} \left[ (cr + 1)^{2H} - 2|cr|^{2H} + (cr - 1)^{2H} \right].$$

(C.45)

Moreover, for all $r \in \mathbb{R}$, $\gamma_Z(r) > 0$.

Proof. For any real lag $r \geq 0$, the autocovariance of fraction Brownian motion can be expressed as

$$\text{cov}(B_H(t + r), B_H(t)) = \frac{1}{2} \left[ (t + r)^{2H} - |r|^{2H} + |t|^{2H} \right].$$

(C.46)

Using Equation (C.46), the lag-$r$ autocovariance function of $Z_H(t)$ is equal to

$$\gamma_Z(r) = \text{cov}(Z_H(t + r), Z_H(t))$$

$$= \text{cov}(B_H(c(t + r)) - B_H(c(t + r) - 1), B_H(ct) - B_H(ct - 1))$$

$$= \text{cov}(B_H(c(t + r)), B_H(ct)) + \text{cov}(B_H(c(t + r) - 1), B_H(ct - 1))$$

$$- \text{cov}(B_H(c(t + r) - 1), B_H(ct)) - \text{cov}(B_H(c(t + r)), B_H(ct - 1))$$

$$= \frac{1}{2} \left[ (cr + 1)^{2H} - 2|cr|^{2H} + (cr - 1)^{2H} \right].$$

(C.47)

Define the function $f(x) = x^{2H}$. We can express

$$|cr + 1|^{2H} - 2|cr|^{2H} + |cr - 1|^{2H} = (cr)^{2H - 2} \frac{f(1 + \frac{1}{cr}) - 2f(1) + f(1 - \frac{1}{cr})}{(cr)^2}$$

(C.48)

Consider $x > 0$ and observe that if $\frac{1}{2} < H < 1$, $f'(x) = 2H x^{2H-1} > 0$ and $f''(x) = 2H (2H - 1)x^{2H-2} > 0$. Hence $f(x)$ is convex. By Jensen’s inequality, $\frac{1}{2} f(1 + \frac{1}{cr}) + \frac{1}{2} f(1 - \frac{1}{cr}) > f(1)$. Therefore, $f(1 + \frac{1}{cr}) - 2f(1) + f(1 - \frac{1}{cr}) > 0$ and $\gamma_Z(r) > 0$. If $r < 0$, it can be shown that $\gamma_Z(r) = \gamma_Z(-r) > 0$. Finally, if $r = 0$, then by Equation (C.48), $\gamma_Z(r) = 1$.

We will use Lemma C.6 for the proof of Lemma C.7.

Lemma C.7. If $d = 0$, $\forall L \in \mathbb{N}$, $-1 < s < 0$, and $-1 < t < 0$,

(i) if $0 < c < \frac{1}{2}$,

$$e^{\gamma_Z(L+s-t)} = \begin{cases} 1 & : \text{if } L > \left\lfloor \frac{1}{c} + 1 \right\rfloor, \\ e^{1-c(L+s-t)} & : \text{if } 1 \leq L \leq \left\lfloor \frac{1}{c} + 1 \right\rfloor. \end{cases}$$
(ii) if $\frac{1}{2} \leq c < 1$

$$e^{\gamma Z(L+s-t)} = \begin{cases} 1 & : \text{if } L > 2, \\ e^{1-c|L+s-t|} & : \text{if } 1 \leq L \leq 2. \end{cases}$$

(iii) if $c \geq 1$

$$e^{\gamma Z(L+s-t)} = \begin{cases} 1 & : \text{if } L \geq 2, \\ e^{1-c|L+s-t|} & : \text{if } L = 1. \end{cases}$$

Proof. If $d = 0$, i.e. $H = \frac{1}{2}$, $\gamma Z(r)$ can be expressed as

$$\gamma Z(r) = \frac{1}{2}c \left( |r + \frac{1}{c}| - 2|r| + |r - \frac{1}{c}| \right).$$

(C.49)

If $|r| \geq \frac{1}{c}$, $\gamma Z(r) = 0$. If $|r| < \frac{1}{c}$, $\gamma Z(r) = 1 - c|r|$. Therefore,

$$e^{\gamma Z(r)} = \begin{cases} 1 & : \text{if } |r| \geq \frac{1}{c}, \\ e^{1-c|r|} & : \text{if } |r| < \frac{1}{c}. \end{cases}$$

Since $-1 \leq s \leq 0$, $-1 \leq t \leq 0$, $|s - t| \leq 1$. $\forall L \in \mathbb{N}$, $0 \leq L + s - t \leq L + 1$. Moreover, if $|L + s - t| \geq \frac{1}{c}$, $e^{\gamma Z(L+s-t)} = 1$. If $|L + s - t| < \frac{1}{c}$, $e^{\gamma Z(L+s-t)} = e^{1-c(L+s-t)}$. Therefore, depending on the range of $c$, $e^{\gamma Z(L+s-t)}$ has different expressions:

(i) if $0 < c < \frac{1}{2}$

$$e^{\gamma Z(L+s-t)} = \begin{cases} 1 & : \text{if } L > \left\lfloor \frac{1}{c} + 1 \right\rfloor, \\ e^{1-c|L+s-t|} & : \text{if } 1 \leq L \leq \left\lfloor \frac{1}{c} + 1 \right\rfloor. \end{cases}$$
(ii) if $\frac{1}{2} \leq c < 1$

\[
\begin{align*}
\gamma_Z(L+s-t) & = \begin{cases} 
1 & : \text{if } L > 2 , \\
\exp\{-c|L+s-t|\} & : \text{if } 1 \leq L \leq 2 .
\end{cases}
\end{align*}
\]

(iii) if $c \geq 1$

\[
\begin{align*}
\gamma_Z(L+s-t) & = \begin{cases} 
1 & : \text{if } L \geq 2 , \\
\exp\{-c|L+s-t|\} & : \text{if } L = 1 .
\end{cases}
\end{align*}
\]

We will use Lemmas C.6, C.7, 2.1 and Corollary 2.2 for the proof of Theorem 3.2.

**Proof of Theorem 3.2**

**Proof.** Assume $d = 0$. Let

\[
h = \begin{cases} 
\lfloor \frac{1}{c} \rfloor + 1 & : \text{if } 0 < c < 1 , \\
\max(\lfloor \frac{1}{c} - 1 \rfloor + 1 , 1) & : \text{if } c \geq 1 ,
\end{cases}
\]

Then $\forall L \in \mathbb{N}$ and $L > h$, $\gamma_Z(L) = 0$ and thus $\text{cov}(A_0, B_L) = \text{cov}(A_L, A_0) = \text{cov}(B_0, B_L) = 0$. From (C.16), if $L > h$, $\text{cov}(r_L, r_0^2) = 0$. Therefore, for all $\tilde{H}m > h$, Equation (3.4) becomes

\[
\text{cov}\left(R_{t,t+\tilde{H}}, RV_{t-\tilde{H},t}\right) = \sum_{L=1}^{h} \text{cov}(r_L, r_0^2)L ,
\]

and (C.37) and (C.39) can be expressed as

\[
\text{var}\left(R_{t,t+\tilde{H}}\right) = (\tilde{H}m)\text{var}(r_0) + 2 \sum_{L=1}^{h} (\tilde{H}m - L)\text{cov}(r_0, r_L) ,
\]

and

\[
\text{var}(RV_{t-\tilde{H},t}) = (\tilde{H}m)\text{var}(r_0^2) + 2 \sum_{L=1}^{h} (\tilde{H}m - L)\text{cov}(r_0^2, r_L^2) .
\]
By Equations (C.51), (C.52), and (C.53), the correlation between $R_{t,t+\tilde{H}}$ and $RV_{t-\tilde{H},t}$ can be expressed as

$$
corr\left(R_{t,t+\tilde{H}}, RV_{t-\tilde{H},t}\right) = \frac{\sum_{L=1}^{h} \text{cov}(r_L, r_0^2) L}{(\tilde{H}m) \sqrt{\left[\text{var}(r_0) + 2 \sum_{L=1}^{h} \left(1 - \frac{L}{Hm}\right) \text{cov}(r_0, r_L)\right] \left[\text{var}(r_0^2) + 2 \sum_{L=1}^{h} \left(1 - \frac{L}{Hm}\right) \text{cov}(r_0^2, r_L^2)\right]}}
$$

(C.54)

Lemma C.6 shows that when $d = 0$, $\forall r \in \mathbb{R}$, $\gamma_Z(r) > 0$. Hence by Equations (2.12) and (C.32), for $L \in \mathbb{N}$, $\text{cov}(r_0, r_L) > 0$ and $\text{cov}(r_0^2, r_L^2) > 0$.

Therefore, for sufficiently large $\tilde{H}$, if $d = 0$,

$$
\text{var}(r_0) + 2 \sum_{L=1}^{h} \left(1 - \frac{L}{Hm}\right) \text{cov}(r_0, r_L) > \text{var}(r_0),
$$

and

$$
\text{var}(r_0^2) + 2 \sum_{L=1}^{h} \left(1 - \frac{L}{Hm}\right) \text{cov}(r_0^2, r_L^2) > \text{var}(r_0^2).
$$

Hence,

$$
\sqrt{\left[\text{var}(r_0) + 2 \sum_{L=1}^{h} \left(1 - \frac{L}{Hm}\right) \text{cov}(r_0, r_L)\right] \left[\text{var}(r_0^2) + 2 \sum_{L=1}^{h} \left(1 - \frac{L}{Hm}\right) \text{cov}(r_0^2, r_L^2)\right]}
$$

is positive and bounded.

From Equation (C.54), as $\tilde{H} \to \infty$, $corr\left(R_{t,t+\tilde{H}}, RV_{t-\tilde{H},t}\right) \to 0$.

\[\square\]

**Lemma C.8.** If $d = 0$, the sample mean $\overline{\Delta N}_n = n^{-1} \sum_{t=1}^{n} \Delta N(t)$ of the counts process is a consistent estimator of $\mu_{\Delta N} = \mathbb{E}[\Delta N(t)]$.

**Proof.** By Equation (A.5) and Lemma C.7 there exists $h \in \mathbb{N}$ such that when $d = 0$, $\forall L > h$, $\text{cov}(\Delta N(L), \Delta N(0)) = 0$. Hence as $n \to \infty$, $\text{var}(\overline{\Delta N}_n) \to 0$. By Chebyshev’s inequality, $\forall \epsilon > 0$, $\mathbb{P}(\overline{\Delta N}_n - \mu_{\Delta N} \geq \epsilon) \leq \frac{\text{var}(\overline{\Delta N}_n)}{\epsilon^2} \to 0$, as $n \to \infty$. Therefore, $\overline{\Delta N}_n \overset{p}{\to} \mu_{\Delta N}$. \[\square\]
Lemma C.9. If $d = 0$, the sample mean $\bar{r}_n$ of the return process is a consistent estimator of $\mu_r = \mathbb{E}[r_t]$.

Proof. By Equation (2.12) and Lemma C.7 there exists $h \in \mathbb{N}$ such that when $d = 0$, $\forall L > h$, $\text{cov}(r_0, r_L) = 0$. Therefore, as $n \to \infty$, $\text{var}(\bar{r}_n) \to 0$. By Chebyshev’s inequality, $\forall \epsilon > 0$, $\mathbb{P}(|\bar{r}_n - \mu_r| \geq \epsilon) \leq \frac{\text{var}(\bar{r}_n)}{\epsilon^2} \to 0$, as $n \to \infty$. Therefore, $\bar{r}_n \xrightarrow{p} \mu_r$.

Lemma C.10. If $d=0$, the sample variance $\hat{\sigma}_r^2$ of the return process is a consistent estimator of $\sigma_r^2$.

Proof. The sample variance $\hat{\sigma}_r^2$ can be expressed as

$$\hat{\sigma}_r^2 = \frac{n}{n-1} \left[ \frac{1}{n} \sum_{t=1}^{n} (r_t - \mu_r)^2 - (\bar{r}_n - \mu_r)^2 \right].$$

From Lemma C.9 if $d = 0$, $\bar{r}_n - \mu_r \xrightarrow{p} 0$. It follows $(\bar{r}_n - \mu_r)^2 \xrightarrow{p} 0$. Moreover, $Z_H(t) = B_H(ct) - B_H(ct - 1)$ is the increment of fractional Brownian Motion. When $d = 0$, there exists $h \in \mathbb{N}$ such that $\forall L > h$, $Z_H(t)$ and $Z_H(t + L)$ are independent. Since the covariance function of $r_t$ is a function of $\gamma Z$, $\forall L > h$, $r_t$ and $r_{t+L}$ are independent. Hence $r_t^2$ and $r_{t+L}^2$ are independent and $\text{cov}(r_t^2, r_{t+L}^2) = 0$ $\forall L \geq h$. Let $Y_t = (r_t - \mu_r)^2$, $\bar{Y}_n = \frac{1}{n} \sum_{t=1}^{n} Y_t$, and $\mathbb{E}[Y_t] = \sigma_r^2$. When $d = 0$, as $n \to \infty$, $\text{var}(\bar{Y}_n) \to 0$. By Chebyshev’s inequality, $\forall \epsilon > 0$, $\mathbb{P}(|\bar{Y}_n - \sigma_r^2| \geq \epsilon) \leq \frac{\text{var}(\bar{Y}_n)}{\epsilon^2} \to 0$. Hence $\bar{Y}_n \xrightarrow{p} \sigma_r^2$. Combining the convergence results of $(\bar{r}_n - \mu_r)^2$ and $\frac{1}{n} \sum_{t=1}^{n} (r_t - \mu_r)^2$, we obtain

$$\frac{n}{n-1} \left[ \frac{1}{n} \sum_{t=1}^{n} (r_t - \mu_r)^2 - (\bar{r}_n - \mu_r)^2 \right] \xrightarrow{p} \sigma_r^2.$$

Lemma C.11. If $d=0$, the sample variance $\hat{\sigma}_{\Delta N}^2$ of the counts process is a consistent estimator of $\sigma_{\Delta N}^2$.

Proof. The sample variance of $\Delta N(t)$ can be expressed as

$$\hat{\sigma}_{\Delta N}^2 = \frac{n}{n-1} \left[ \frac{1}{n} \sum_{t=1}^{n} (\Delta N(t) - \mu_{\Delta N})^2 - (\Delta N_n - \mu_{\Delta N})^2 \right].$$
From Lemma C.8 if \( d = 0 \), \( \Delta N_n - \mu_{\Delta N} \overset{p}{\to} 0 \). It follows \((\Delta N_n - \mu_{\Delta N})^2 \overset{p}{\to} 0 \). Since the covariance function of \( \Delta N \) is a function of \( \gamma_Z \), if \( d = 0 \), there exists \( L \geq h \) such that \( \Delta N(t) \) and \( \Delta N(t+L) \) are independent. Hence \((\Delta N(t))^2 \) and \((\Delta N(t+L))^2 \) are independent. Let \( \omega_t = (\Delta N(t) - \mu_{\Delta N})^2 \), \( \bar{\omega}_n = \frac{1}{n} \sum_{t=1}^{n}(\Delta N(t) - \mu_{\Delta N})^2 \), and \( \mu_\omega = \mathbb{E}[\omega_t] = \sigma^2_{\Delta N} \). When \( d = 0 \), as \( n \to \infty \), \( \text{var}(\bar{\omega}_n) \to 0 \). By Chebyshev’s inequality, \( \forall \epsilon > 0 \), \( \mathbb{P}(|\bar{\omega}_n - \sigma^2_{\Delta N}| \geq \epsilon) \leq \frac{\text{var}(\bar{\omega}_n)}{\epsilon^2} \to 0 \). Hence \( \bar{\omega}_n \overset{p}{\to} \sigma^2_{\Delta N} \). Combining the convergence results of \((\Delta N_n - \mu_{\Delta N})^2 \) and \( \frac{1}{n} \sum_{t=1}^{n}(\Delta N(t) - \mu_{\Delta N})^2 \), we obtain

\[
\frac{n}{n-1} \left[ \frac{1}{n} \sum_{t=1}^{n}(\Delta N(t) - \mu_{\Delta N})^2 - (\Delta N_n - \mu_{\Delta N})^2 \right] \overset{p}{\to} \sigma^2_{\Delta N}.
\]

\[\square\]

**Lemma C.12.** If \( d=0 \), the sample lag-\( L \) autocovariance \( \hat{\gamma}_{\Delta N}(L) \) of the counts is a consistent estimator of \( \gamma_{\Delta N}(L) \).

**Proof.** By definition the sample lag-\( L \) autocovariance of \( \Delta N(t) \) is

\[
\hat{\gamma}_{\Delta N}(L) = \frac{1}{n} \sum_{t=1}^{n-L} [(\Delta N(t) - \Delta N(t+L)) \Delta N(t+L) - \Delta N(t+L)]
\]

where \( \Delta N_{1,n-L} = \frac{1}{n-L} \sum_{t=1}^{n-L} \Delta N(t) \) and \( \Delta N_{L+1,n} = \frac{1}{n-L} \sum_{t=L+1}^{n} \Delta N(t) \). Hence \( \hat{\gamma}_{\Delta N}(L) \) can be expressed as

\[
\hat{\gamma}_{\Delta N}(L) = \frac{n-L}{n} \left\{ \frac{1}{n-L} \sum_{t=1}^{n-L} [((\Delta N(t) - \mu_{\Delta N})(\Delta N(t+L) - \mu_{\Delta N})] 
- \frac{1}{n-L} (\Delta N_{1,n-L} - \mu_{\Delta N}) \sum_{t=1}^{n-L} (\Delta N(t+L) - \mu_{\Delta N})
- \frac{1}{n-L} (\Delta N_{L+1,n} - \mu_{\Delta N}) \sum_{t=1}^{n-L} (\Delta N(t) - \mu_{\Delta N})
+ (\Delta N_{1,n-L} - \mu_{\Delta N})(\Delta N_{L+1,n} - \mu_{\Delta N}) \right\}
\]

\[
= \frac{n-L}{n} \left\{ \frac{1}{n-L} \sum_{t=1}^{n-L} [((\Delta N(t) - \mu_{\Delta N})(\Delta N(t+L) - \mu_{\Delta N})]
- (\Delta N_{1,n-L} - \mu_{\Delta N})(\Delta N_{L+1,n} - \mu_{\Delta N}) \right\}.
\]

By Lemma C.8, \( \Delta N(t) \overset{p}{\to} \mu_{\Delta N} \) and \( \Delta N(t+L) \overset{p}{\to} \mu_{\Delta N} \). Hence \( (\Delta N(t) - \mu_{\Delta N})(\Delta N(t+L) - \mu_{\Delta N}) \overset{p}{\to} 0 \).
Let $\hat{W}_t = (\Delta N(t) - \mu_{\Delta N})(\Delta N(t + L) - \mu_{\Delta N})$ and $\mu_{\hat{W}} = \mathbb{E}\hat{W}_t = \gamma_{\Delta N}(L)$. Define $\hat{W}_n = \frac{1}{n-\hat{n}} \sum_{t=1}^{n-\hat{n}} \hat{W}_t$. When $d = 0$, there exists $L > h$ such that $\Delta N(t) \Delta N(t + L)$ and $\Delta N(t + i) \Delta N(t + i + L)$ are independent $\forall \ i > L + h$. Hence $\text{cov}(\hat{W}_t, \hat{W}_{t+i}) = 0 \ \forall \ i > L + h$. When $d = 0$, as $n \to \infty$, $\text{var}(\hat{W}_n) \to 0$. By Chebyshev’s inequality, $\mathbb{P}(|\hat{W}_n - \mu_{\hat{W}}| \geq \epsilon) \leq \frac{\text{var}(\hat{W}_n)}{\epsilon^2}$, as $\epsilon \to 0$. Combining the convergence results of $(\Delta N(t) - \mu_{\Delta N})(\Delta N(t + L) - \mu_{\Delta N})$ and $\hat{W}_n$ under $0 < L \leq \hat{H}$, we obtain $\hat{\gamma}_{\Delta N}(L) \xrightarrow{p} \gamma_{\Delta N}(L)$.

We will use Lemmas $\text{C.8}$, $\text{C.9}$, $\text{C.10}$, $\text{C.11}$ and $\text{C.12}$ for the proof of Theorem $\text{A.2}$.

**Proof of Theorem $\text{A.2}$**

*Proof.* By Lemmas $\text{C.9}$ and $\text{C.8}$, $\bar{r}_n \xrightarrow{p} \mu_r$ and $\overline{\Delta N}_n \xrightarrow{p} \mu_{\Delta N}$. Hence the joint vector $(\bar{r}_n, \overline{\Delta N}_n) \xrightarrow{p} (\mu_r, \mu_{\Delta N})$.

Define $g(\bar{r}_n, \overline{\Delta N}_n) = \frac{\bar{r}_n}{\overline{\Delta N}_n}$. It follows that $g(\bar{r}_n, \overline{\Delta N}_n) \xrightarrow{p} g(\mu_r, \mu_{\Delta N})$. Therefore, $\hat{\mu} \xrightarrow{p} \mu$.

Since $\hat{\lambda} = \frac{\overline{\Delta N}_n}{c^2}$, by Lemma $\text{C.8}$, $\hat{\lambda} \xrightarrow{p} \frac{\mu_{\Delta N}}{c^2} = \lambda$.

By Lemmas $\text{C.10}$ and $\text{C.11}$, $\hat{\sigma}_r^2 \xrightarrow{p} \sigma_r^2$ and $\hat{\sigma}_{\Delta N}^2 \xrightarrow{p} \sigma_{\Delta N}^2$. From $\text{A.15}$, $\hat{\sigma}_r^2$ is a function of $\hat{\sigma}_r^2$, $\hat{\sigma}_{\Delta N}^2$, $\hat{\mu}$, and $\hat{\lambda}$. It follows that $\hat{\sigma}_r^2 \xrightarrow{p} \sigma_r^2$.

We substitute $\hat{\lambda}$, $d = 0$ and $L = 1$ into Equation $\text{A.5}$ to evaluate $\gamma_{\Delta N}(1)$ under $d = 0$ and estimate $c$ by solving the equation

$$\gamma_{\Delta N}(1) - \hat{\gamma}_{\Delta N}(1) = 0,$$

(C.56)

where $\hat{\gamma}_{\Delta N}(1)$ is the sample lag-one autocovariance of the count. Therefore, we can express $\hat{c}$ as a function of $\hat{\lambda}$ and $\hat{\gamma}_{\Delta N}(1)$. Let $\hat{c} = g\hat{c}(\hat{\lambda}, \hat{\gamma}_{\Delta N}(1))$ and $c = g\hat{c}(\lambda, \gamma_{\Delta N}(1))$. From Lemma $\text{C.12}$, $\hat{\gamma}_{\Delta N}(1) \xrightarrow{p} \gamma_{\Delta N}(1)$. Using the fact $\hat{\lambda} \xrightarrow{p} \lambda$, we obtain $g\hat{c}(\hat{\lambda}, \hat{\gamma}_{\Delta N}(1)) \xrightarrow{p} g\hat{c}(\lambda, \gamma_{\Delta N}(1))$.

*Lemma C.13.* Let $\gamma_R(L) = \text{cov}(R_{t,i+\hat{H}}, R_{t+L,i+L+\hat{H}})$ and $\gamma_r(k) = \text{cov}(r_t, r_{t+k})$. If $d = 0$, $\forall \ 0 < L \leq \hat{H}$,
i) if $\tilde{H} - L$ is odd,

$$
\gamma_R(L) = [(\tilde{H} - L)m] \left[ \text{var}(r_t) + 2 \sum_{k=1}^{m} \gamma_r(k) \right] + 2 \sum_{k=m+1}^{(\tilde{H}+L)m} [(\tilde{H} - L)m - k] \gamma_r(k) + \sum_{k=(\tilde{H}-L)m+1}^{(\tilde{H}+L)m-1} [(\tilde{H} + L)m - k] \gamma_r(k). 
$$

(C.57)

\[ (C.57) \]

\[ (C.57) \]

ii) if $\tilde{H} - L$ is even,

$$
\gamma_R(L) = [(\tilde{H} - L)m] \left[ \text{var}(r_t) + 2 \sum_{k=1}^{(\tilde{H}-L)m} \gamma_r(k) \right] + \sum_{k=(\tilde{H}-L)m+1}^{(\tilde{H}+L)m-1} [(\tilde{H} + L)m - k] \gamma_r(k).
$$

(C.58)

It follows $\gamma_R(L) = O(\tilde{H})$.

Proof.

$$
\text{cov}(R_{t,i+\tilde{H}}, R_{t+i+L,i+L+\tilde{H}}) = \sum_{j=lm+1}^{(i+\tilde{H})m} \sum_{i=lm+1}^{(i+\tilde{H})m} \text{cov}(r_j, r_{i+Lm})
$$

$$
= \sum_{j=1}^{Hm} \sum_{i=1}^{Hm} \text{cov}(r_j, r_{i+Lm})
$$

(C.59)

Define $k = |i + Lm - j|$. If $\tilde{H} - L$ is odd, then (C.59) has the expression of (C.57). If $\tilde{H} - L$ is even, then (C.59) has the expression of (C.58).

Corollary C.14. If $L = \tilde{H}$, then $\gamma_R(\tilde{H}) = \sum_{k=1}^{2\tilde{H}m-1} [2\tilde{H}m - k] \gamma_r(k)$. When $d = 0$, $r_t$ is $h$-dependent, and thus $\gamma_R(\tilde{H}) = \sum_{k=1}^{h} [2\tilde{H}m - k] \gamma_r(k)$. Hence $R_{t,i+\tilde{H}}$ is $\tilde{H} + h/m$ dependent.

Lemma C.15. If $d = 0$,

$$
\text{var} \left( R_{t,i+\tilde{H}}^2 \right) \sim \tilde{H}^3 C^*,
$$

$$
\text{var} \left( \tilde{R}_{t,i+\tilde{H}}^2 \right) \sim \tilde{H}^3 C^{**}
$$

where $C^*$ and $C^{**}$ are positive constants.
Proof. Let \( R_{t,\tilde{t}+H}^* = R_{t,\tilde{t}+H} - \mathbb{E}[R_{t,\tilde{t}+H}] = \sum_{i=tm+1}^{(i+\tilde{H})m} (r_i - \mathbb{E}[r_i]) \). Let \( r_t^* = r_t - \mathbb{E}[r_t] \). Then
\[
R_{t,\tilde{t}+H}^2 = \sum_{i=tm+1}^{(i+\tilde{H})m} \sum_{j=tm+1}^{(i+\tilde{H})m} r_i^* r_j^*.
\]
Hence
\[
\text{var}(R_{t,\tilde{t}+H}^2) = \text{var}(R_{t,\tilde{t}+H}^2) + 4\mathbb{E}[R_{t,\tilde{t}+H}]\text{cov}(R_{t,\tilde{t}+H}, R_{t,\tilde{t}+H}^2) + 4(\mathbb{E}[R_{t,\tilde{t}+H}])^2 \text{var}(R_{t,\tilde{t}+H})\,.
\]

Applying Proposition 3.2.1 by Peccati and Taqqu (2011), we obtain
\[
\text{var}(R_{t,\tilde{t}+H}^2) = \text{cov}(R_{t,\tilde{t}+H}^2, R_{t,\tilde{t}+H}^2)
\]
\[
= \sum_{i=tm+1}^{(i+\tilde{H})m} \sum_{j=tm+1}^{(i+\tilde{H})m} \sum_{u=tm+1}^{(i+\tilde{H})m} \sum_{v=tm+1}^{(i+\tilde{H})m} \text{cov}(r_i^* r_j^*, r_u^* r_v^*)
\]
\[
= \sum_{i=tm+1}^{(i+\tilde{H})m} \sum_{j=tm+1}^{(i+\tilde{H})m} \sum_{u=tm+1}^{(i+\tilde{H})m} \sum_{v=tm+1}^{(i+\tilde{H})m} \text{cum}(r_i^*, r_j^*, r_u^*, r_v^*)
\]
\[
+ 2 \sum_{i=tm+1}^{(i+\tilde{H})m} \sum_{j=tm+1}^{(i+\tilde{H})m} \sum_{u=tm+1}^{(i+\tilde{H})m} \sum_{v=tm+1}^{(i+\tilde{H})m} \text{cov}(r_i^*, r_u^*) \text{cov}(r_j^*, r_v^*)
\]
\[
= \sum_{i=tm+1}^{(i+\tilde{H})m} \sum_{j=tm+1}^{(i+\tilde{H})m} \sum_{u=tm+1}^{(i+\tilde{H})m} \sum_{v=tm+1}^{(i+\tilde{H})m} \text{cum}(r_i, r_j, r_u, r_v)
\]
\[
+ 2 \sum_{i=tm+1}^{(i+\tilde{H})m} \sum_{j=tm+1}^{(i+\tilde{H})m} \sum_{u=tm+1}^{(i+\tilde{H})m} \sum_{v=tm+1}^{(i+\tilde{H})m} \text{cov}(r_i, r_v) \text{cov}(r_u, r_j)
\]

When \( d = 0 \), \( \{r_i\} \) is \( h \)-dependent. For a fixed \( i \), if \( |i - v| > h \), then \( r_i \) is independent of \( r_v \). Similarly, for a fixed \( u \), if \( |u - j| > h \), then \( r_u \) is independent of \( r_j \).

Hence
\[
\text{cov} = \sum_{i=tm+1}^{(i+\tilde{H})m} \text{cov}(r_i, r_v) \sum_{u=tm+1}^{(i+\tilde{H})m} \text{cov}(r_u, r_j)
\]
\[
= \sum_{i=tm+1}^{(i+\tilde{H})m} \text{cov}(r_i, r_v) \sum_{u=tm+1}^{(i+\tilde{H})m} \text{cov}(r_u, r_j)
\]
\[
\sim \tilde{H}^2 m \sum_{|i-v| \leq h} \sum_{|u-j| \leq h} \text{cov}(r_i, r_v) \text{cov}(r_u, r_j).
\]

If any group of \( r_i, r_j, r_u, r_v \) is independent of the remaining elements, \( \text{cum}(r_i, r_j, r_u, r_v) = 0 \). Hence for fixed \( i \), if \( |i - v| \leq h \), \( |i - j| \leq h \), and \( |i - u| \leq h \), \( \text{cum}(r_i, r_j, r_u, r_v) \neq 0 \). Thus
\[
\text{cum} \sim \tilde{H} m \sum_{|i-j| \leq h} \sum_{|i-u| \leq h} \sum_{|i-v| \leq h} \text{cum}(r_i, r_j, r_u, r_v).
\]
Combining these results, \( \text{var}(R_{t,i+H}^*) \sim \tilde{H}^2 C_1 \), where

\[
C_1 = m \sum_{|i-v| \leq h} \sum_{|u-j| \leq h} \text{cov}(r_i, r_v) \text{cov}(r_u, r_j) \\
+ m \sum_{|i-j| \leq h} \sum_{|i-u| \leq h} \sum_{|i-v| \leq h} \text{cum}(r_i, r_j, r_u, r_v).
\]

Similarly,

\[
\text{cov} \left( R_{t,i+H}, R_{t,i+H}^2 \right) = \left( \frac{(i+	ilde{H})m}{(i+	ilde{H})m} \right) \sum_{i=tm+1}^{(i+	ilde{H})m} \sum_{j=tm+1}^{(i+	ilde{H})m} \sum_{u=tm+1}^{(i+	ilde{H})m} \text{cov}(r_i, r_j, r_u) \\
= \sum_{i=tm+1}^{(i+	ilde{H})m} \sum_{j=tm+1}^{(i+	ilde{H})m} \sum_{u=tm+1}^{(i+	ilde{H})m} \text{cum}(r_i, r_j, r_u) \\
+ 2 \sum_{i=tm+1}^{(i+	ilde{H})m} \sum_{j=tm+1}^{(i+	ilde{H})m} \sum_{u=tm+1}^{(i+	ilde{H})m} \mathbb{E}[r_v] \text{cov}(r_i, r_j).
\]

Applying the similar argument as for \( \text{cum}_1 \) and \( \text{cov}_1 \), we obtain

\[
\text{cum}_2 \sim \tilde{H}m \sum_{|i-u| \leq h} \sum_{|i-v| \leq h} \text{cum}(r_i, r_j, r_u) \\
\text{cov}_2 = \tilde{H}m \mathbb{E}[r_i] \sum_{i=tm+1}^{(i+	ilde{H})m} \sum_{j=tm+1}^{(i+	ilde{H})m} \text{cov}(r_i, r_j) \mathbb{1}\{|i - j| \leq h\} \\
= \tilde{H}m \mathbb{E}[r_i] \sum_{j=tm+1}^{(i+	ilde{H})m} \left( (i+	ilde{H})m - |j - h| \right) \text{cov}(r_0, r_{j-\tilde{H}-h}) \\
\sim \tilde{H}^2 C_2,
\]

where \( C_2 \) is a positive constant. Hence \( \text{cov} \left( R_{t,i+H}, R_{t,i+H}^2 \right) \sim \tilde{H}^2 C_2 \). Combining the results for \( \text{var}(R_{t,i+H}^*) \) and \( \text{cov} \left( R_{t,i+H}, R_{t,i+H}^2 \right) \) as well as the results that \( \mathbb{E}[R_{t,i+H}] = \tilde{H}m \mathbb{E}[r_i] \) and \( \text{var}(R_{t,i+H}) \sim \tilde{H}C_3 \) when \( d = 0 \), where \( C_3 \) is a constant (see (C.39)), it is straightforward to show that \( \text{var}(R_{t,i+H}^2) \sim \tilde{H}^3 C^* \), where \( C^* \) is a constant. With similar arguments, it can be shown \( \text{var}(\tilde{R}_{t,i+H}^2) \sim \tilde{H}^3 C^{**} \).
Lemma C.16. If \( d = 0 \),
\[
\text{var}\left( RV_{i-h,i}^2 \right) \sim \tilde{H}^3 D^*
\]
where \( D^* \) is a positive constant.

Proof. We skip the proof since it is similar to the proof of C.15. \( \square \)

Corollary C.17. For every \( 0 < L \leq \tilde{H} \),
\[
\text{cov}\left( R_{i-h,i}^2, R_{i+L,h}^2 \right) \sim D_1 \tilde{H}^3
\]
\[
\text{cov}\left( RV_{i-h,i}^2, RV_{i+L-h,i+L}^2 \right) \sim D_2 \tilde{H}^3
\]
where \( D_1 \) and \( D_2 \) are positive constants.

Lemma C.18.
\[
\text{var}(r_t^2) = \mathbb{E}[r_t^4] - (\text{var}(r_t))^2 - 2(\mathbb{E}[r_t])^2\text{var}(r_t) - (\mathbb{E}[r_t])^4 \quad (C.61)
\]
where \( \text{var}(r_t) \) is given by (2.13),
\[
\mathbb{E}[r_t^4] = (\mu^4 + 6\mu^2\sigma_e^2 + 3\sigma_e^4)\mathbb{E}[\nu] + (7\mu^4 + 18\mu^2\sigma_e^2)\mathbb{E}[\nu^2] + (6\mu^4 + 6\mu^2\sigma_e^2)\mathbb{E}[\nu^3] + \mu^4\mathbb{E}[\nu^4] \quad , \quad (C.62)
\]
\( \nu = \mathbb{E}[\Delta N(t)|\Lambda] \), and \( \mathbb{E}[\nu], \mathbb{E}[\nu^2], \mathbb{E}[\nu^3], \text{and } \mathbb{E}[\nu^4] \) are given by (C.73), (C.74), (C.75), and (C.76).

Proof. Applying Proposition 3.2.1 by Peccati and Taqqu(2011), we can express \( \text{var}(r_t^2) \) as
\[
\text{var}(r_t^2) = \text{cum}(r_t, r_t, r_t, r_t) + 4\mathbb{E}[r_t]\text{cum}(r_t, r_t, r_t) + 2\text{cum}(r_t, r_t)\text{cum}(r_t, r_t) + 4(\mathbb{E}[r_t])^2\text{cum}(r_t, r_t) \quad (C.63)
\]
Let \( r_t^* = r_t - \mathbb{E}[r_t] \). Then \( \text{cum}(r_t, r_t, r_t, r_t) = \text{cum}(r_t^*, r_t^*, r_t^*, r_t^*) \), and \( \text{cum}(r_t, r_t, r_t) = \text{cum}(r_t^*, r_t^*, r_t^*) \). Therefore,
\[
\text{var}(r_t^2) = \text{cum}(r_t^*, r_t^*, r_t^*, r_t^*) + 4\mathbb{E}[r_t]\text{cum}(r_t^*, r_t^*, r_t^*) + 2(\text{var}(r_t))^2 + 4(\mathbb{E}[r_t])^2\text{var}(r_t) \quad (C.64)
\]
where
\[
\text{cum}(r_t^*, r_t^*, r_t^*) = -3\mathbb{E}[(r_t^*)^2] + \mathbb{E}[(r_t^*)^4]
\]
\[
\text{cum}(r_t^*, r_t^*, r_t^*) = \mathbb{E}[(r_t^*)^3]
\]
Let $\mu_r = \mathbb{E}[r_t]$. Then

\begin{align*}
\mathbb{E}[(r_t^*)^4] &= \mathbb{E}[r_t^4] - 4\mu_r\mathbb{E}[r_t^3] + 6\mu_r^2\mathbb{E}[r_t^2] - 3\mu_r^4 \quad (C.65) \\
\mathbb{E}[(r_t^*)^3] &= \mathbb{E}[r_t^3] - 3\mu_r\mathbb{E}[r_t^2] + 2\mu_r^3 \quad (C.66) \\
\mathbb{E}[(r_t^*)^2] &= \text{var}(r_t) \quad (C.67)
\end{align*}

Plugging (C.65), (C.66), and (C.67) into (C.64), we obtain

\begin{align*}
\text{var}(r_t^2) &= -(\text{var}(r_t))^2 + \mathbb{E}[r_t^4] - 2\mu_r^2\text{var}(r_t) - \mu_r^4 \quad (C.68)
\end{align*}

where

\begin{align*}
\mathbb{E}[r_t^4] &= \mathbb{E} \left[ \mu \Delta N(t) + \sum_{k=N(t-1)+1}^{N(t)} e_k \right]^4 \\
&= \mu^4 \mathbb{E}[(\Delta N(t))^4] + 6\mu^2\sigma_e^2\mathbb{E}[(\Delta N(t))^3] + 3\mathbb{E}[\Delta N(t)]\sigma_e^4 \quad (C.69)
\end{align*}

Conditional on $\Lambda$, $\Delta N(t)$ is Poisson process. Let $\nu = \mathbb{E}[\Delta N(t)|\Lambda]$. $\nu = \lambda \int_{t-1}^t e^{Z_H(s)} ds$. Hence

\begin{align*}
\mathbb{E}[(\Delta N(t))^4] &= \mathbb{E}[\mathbb{E}[(\Delta N(t))^4|\Lambda]] = \mathbb{E}[\nu] + 7\mathbb{E}[\nu^2] + 6\mathbb{E}[\nu^3] + \mathbb{E}[\nu^4] \quad (C.70) \\
\mathbb{E}[(\Delta N(t))^3] &= \mathbb{E}[\mathbb{E}[(\Delta N(t))^3|\Lambda]] = \mathbb{E}[\nu] + 3\mathbb{E}[\nu^2] + \mathbb{E}[\nu^3] \quad (C.71)
\end{align*}

and thus $\mathbb{E}[r_0^4]$ can be represented as

\begin{align*}
\mathbb{E}[r_t^4] &= (\mu^4 + 6\mu^2\sigma_e^2 + 3\sigma_e^4)\mathbb{E}[\nu] + (7\mu^4 + 18\mu^2\sigma_e^2)\mathbb{E}[\nu^2] + (6\mu^4 + 6\mu^2\sigma_e^2)\mathbb{E}[\nu^3] + \mu^4\mathbb{E}[\nu^4] \quad (C.72)
\end{align*}
where

\[ E[\nu] = \lambda e^{1/2} \quad (C.73) \]

\[ E[\nu^2] = \lambda^2 \int_{-1}^{0} \int_{-1}^{0} \mathbb{E}[e^{Z_H(u)} e^{Z_H(v)}] dudv = \lambda^2 \int_{-1}^{0} \int_{-1}^{0} e^{1 + \gamma_2(t - s)} dt ds \quad (C.74) \]

\[ E[\nu^3] = \lambda^3 \int_{-1}^{0} \int_{-1}^{0} \int_{-1}^{0} \mathbb{E}[e^{Z_H(t)} e^{Z_H(s)} e^{Z_H(u)}] dt ds du \]

\[ = \lambda^3 \int_{-1}^{0} \int_{-1}^{0} \int_{-1}^{0} e^{3 + \gamma_2(t - s) + \gamma_2(t - u) + \gamma_2(s - u)} dt ds du \quad (C.75) \]

\[ E[\nu^4] = \lambda^4 \int_{-1}^{0} \int_{-1}^{0} \int_{-1}^{0} \int_{-1}^{0} \mathbb{E}[e^{Z_H(t)} e^{Z_H(s)} e^{Z_H(u)} e^{Z_H(v)}] dt ds du dv \]

\[ = \lambda^4 \int_{-1}^{0} \int_{-1}^{0} \int_{-1}^{0} \int_{-1}^{0} e^{2 + \gamma_2(t - s) + \gamma_2(t - u) + \gamma_2(v - u) + \gamma_2(s - v) + \gamma_2(u - v)} dt ds du dv \quad (C.76) \]

**Lemma C.19.** Define

\[ \tilde{R}_{\tilde{H}} = \frac{1}{\tilde{T} - 2\tilde{H}} \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} R_{i,\tilde{i}+\tilde{H}} \]

\[ \tilde{R}V_{\tilde{H}} = \frac{1}{\tilde{T} - 2\tilde{H}} \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} RV_{\tilde{i} - \tilde{H},\tilde{i}} , \]

where \( \tilde{H} = \tilde{T}^\kappa, \) \( 0 < \kappa < 1. \) If \( d = 0, \) for every \( \epsilon' > 0, \)

\[ \frac{\tilde{R}_{\tilde{H}}}{\mathbb{E}[R_{i,\tilde{i}+\tilde{H}}]} = 1 + o_p(\tilde{T}^{-1/2 + \epsilon'/2}) , \]

\[ \frac{\tilde{R}V_{\tilde{H}}}{\mathbb{E}[RV_{\tilde{i} - \tilde{H},\tilde{i}}]} = 1 + o_p(\tilde{T}^{-1/2 + \epsilon'/2}) . \]

**Proof.** The expected value of \( \frac{\tilde{R}_{\tilde{H}}}{\mathbb{E}[R_{i,\tilde{i}+\tilde{H}}]} \) is equal to

\[ \mathbb{E} \left[ \frac{\tilde{R}_{\tilde{H}}}{\mathbb{E}[R_{i,\tilde{i}+\tilde{H}}]} \right] = 1 , \]

and its variance is equal to

\[ \text{var} \left( \frac{\tilde{R}_{\tilde{H}}}{\mathbb{E}[R_{i,\tilde{i}+\tilde{H}}]} \right) = \frac{\text{var} \left( \tilde{R}_{\tilde{H}} \right)}{(\mathbb{E}[R_{i,\tilde{i}+\tilde{H}}])^2} = \frac{1}{m^2(\mathbb{E}[r_i])^2} \left[ \frac{\text{var} \left( \tilde{R}_{\tilde{H}} \right)}{\tilde{H}^2} \right] , \]
where
\[
\text{var} \left( \frac{\bar{R}_H}{\mathbb{E}[R_{t, \hat{H}}]} \right) = \frac{1}{(\bar{T} - 2\bar{H})^2} \left[ \text{var} \left( \frac{\bar{T} - \hat{H}}{\sum_{t=H}^{\bar{T} - \hat{H}} R_{t, \hat{H}}} \right) \right],
\]
and
\[
\text{var} \left( \sum_{t=H}^{\bar{T} - \hat{H}} R_{t, \hat{H}} \right) = (\bar{T} - 2\bar{H}) \text{var}(R_{t, \hat{H}}) + \sum_{t=H+1}^{\bar{T} - \hat{H}} \sum_{\hat{H}+1}^{\bar{T} - \hat{H}} \text{cov} \left( R_{t, \hat{H}}, R_{k, \hat{H}} \right)
\]
\[
= (\bar{T} - 2\bar{H}) \text{var}(R_{t, \hat{H}}) + 2 \sum_{0 < L < \bar{T} - 2\bar{H}} (\bar{T} - 2\bar{H} - L) \text{cov} \left( R_{t, \hat{H}}, R_{t+L, \hat{H}+L} \right)
\]

Let \( \gamma_R(L) = \text{cov}(R_{t, \hat{H}}, R_{t+L, \hat{H}+L}) \).

From Corollary C.14, if \( d = 0, R_{t, \hat{H}} \) is \( \hat{H} + h/m \) dependent. Hence we can express

\[
\sum_{0 < L < \bar{T} - 2\bar{H}} (\bar{T} - 2\bar{H} - L) \text{cov} \left( R_{t, \hat{H}}, R_{t+L, \hat{H}+L} \right) = \sum_{L=1}^{(\hat{H} + h/m) \land (\bar{T} - 2\bar{H})} (\bar{T} - 2\bar{H} - L) \gamma_R(L).
\]

By Lemmas C.5 and C.13, \( \text{var}(R_{t, \hat{H}}) \sim C_{\mu, \lambda} \hat{H}, \) and \( \forall 0 < L \leq \hat{H}, \gamma_R(L) \sim C_4 \hat{H} \), where both \( C_{\mu, \lambda} \) and \( C_4 \) are positive constants.

If \( \hat{H} + h/m < \bar{T} - 2\bar{H} \),

\[
\text{var} \left( \sum_{t=H}^{\bar{T} - \hat{H}} R_{t, \hat{H}} \right) = O((\bar{T} - 2\bar{H}) \bar{H}^2)
\]
and thus

\[
\text{var} \left( \frac{\bar{R}_H}{\mathbb{E}[R_{t, \hat{H}}]} \right) = O \left( \frac{\hat{H}^2}{\bar{T}} \right).
\]

It follows that \( \text{var} \left( \frac{\bar{R}_H}{\mathbb{E}[R_{t, \hat{H}}]} \right) = O(\bar{T}^{-1}) \). Note that \( \hat{H} + h/m < \bar{T} - 2\bar{H} \) implies \( 0 < \kappa < \frac{\log(\bar{T} - h/m) - \log^3}{\log \bar{T}} \), which converges to 1 as \( \bar{T} \to \infty \).

By Chebyshev’s inequality, \( \forall \epsilon' > 0 \), we obtain

\[
P \left( \left| \frac{\bar{R}_H}{\mathbb{E}[R_{t, \hat{H}}]} - 1 \right| \geq \epsilon \right) \leq \frac{(\bar{T}^{-1}\epsilon') \text{var} \left( \frac{\bar{R}_H}{\mathbb{E}[R_{t, \hat{H}}]} \right)}{\epsilon'^2} = O(\bar{T}^{-\epsilon'}) \to 0,
\]

62
as $\bar{T} \to \infty$.

The proof of $RV_{\tilde{t} - \tilde{H}, \tilde{t}}$ follows from a similar argument.

**Corollary C.20.** If $d = 0$, for $0 < \kappa < 1/2$,

$$
\frac{\tilde{R}_{\tilde{H}} - \mathbb{E}[R_{\tilde{t}, \tilde{t} + \tilde{H}}]}{RV_{\tilde{H}} - \mathbb{E}[RV_{\tilde{t}, \tilde{t} + \tilde{H}}]} \xrightarrow{p} 0
$$

**Remark C.21.** Follow the proof of Lemma C.19 and Corollary C.20, the same limiting results apply to $\tilde{R}_{\tilde{H}}$.

**Lemma C.22.** When $d = 0$, for $0 < \kappa < 1$ and for every $\epsilon' > 0$,

$$
\frac{1}{\bar{T} - 2\bar{H}} \sum_{i=\bar{H}+1}^{\bar{T} - \bar{H}} \frac{R_{\tilde{t}, \tilde{t} + \tilde{H}}}{\mathbb{E}[R_{\tilde{t}, \tilde{t} + \tilde{H}}]} = 1 + o_p(\bar{T}^{-1/2+\epsilon'/2})
$$

**Proof.** The variance of the ratio is equal to

$$
\text{var} \left( \frac{1}{\bar{T} - 2\bar{H}} \sum_{i=\bar{H}+1}^{\bar{T} - \bar{H}} \frac{R_{\tilde{t}, \tilde{t} + \tilde{H}}}{\mathbb{E}[R_{\tilde{t}, \tilde{t} + \tilde{H}}]} \right) = \frac{1}{(\mathbb{E}[R_{\tilde{t}, \tilde{t} + \tilde{H}}])^2(\bar{T} - 2\bar{H})^2} \text{var} \left( \sum_{i=\bar{H}+1}^{\bar{T} - \bar{H}} R_{\tilde{t}, \tilde{t} + \tilde{H}} \right)
$$

Note that

$$
\mathbb{E}[R_{\tilde{t}, \tilde{t} + \tilde{H}}] = \text{var}(R_{\tilde{t}, \tilde{t} + \tilde{H}}) + (\mathbb{E}[R_{\tilde{t}, \tilde{t} + \tilde{H}}])^2
$$

Since var($R_{\tilde{t}, \tilde{t} + \tilde{H}}$) $\sim C_{\lambda, \mu} \tilde{H}$ and $\mathbb{E}(R_{\tilde{t}, \tilde{t} + \tilde{H}})$ is proportional to $\tilde{H}$, we have $\mathbb{E}[R_{\tilde{t}, \tilde{t} + \tilde{H}}] = O(\tilde{H}^2)$. From Lemma C.15, we obtain

$$
\sum_{i=\bar{H}+1}^{\bar{T} - \bar{H}} \text{var} \left( R_{\tilde{t}, \tilde{t} + \tilde{H}} \right) = O((\bar{T} - 2\bar{H})\tilde{H}^3). \quad (C.78)
$$

By Corollary C.17, if $d = 0$, for every $0 < L \leq \tilde{H}$, $\text{cov}(R_{\tilde{t}, \tilde{t} + \tilde{H}}, R_{\tilde{t} + L, \tilde{t} + L + \tilde{H}}) \sim D_1 \tilde{H}^3$. Hence, if $\tilde{H} + h/m < \bar{T} - 2\tilde{H}$, which implies that $0 < \kappa < 1$,

$$
\sum_{L=1}^{\tilde{H} + h/m \wedge (\bar{T} - 2\tilde{H})} (\bar{T} - 2\tilde{H} - L) \text{cov} \left( R_{\tilde{t}, \tilde{t} + \tilde{H}}, R_{\tilde{t} + L, \tilde{t} + L + \tilde{H}} \right) = O((\bar{T} - 2\tilde{H})\tilde{H}^4). \quad (C.79)
$$
Combining (C.78) and (C.79), we obtain

\[
\text{var} \left( \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} R^2_{i,\tilde{t}+\tilde{H}} \right) = O((\tilde{T} - 2\tilde{H})\tilde{H}^4),
\]

and thus

\[
\text{var} \left( \frac{1}{\tilde{T} - 2\tilde{H}} \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} \frac{R^2_{i,\tilde{t}+\tilde{H}}}{\mathbb{E}[R^2_{i,\tilde{t}+\tilde{H}}]} \right) = O(\tilde{T}^{-1}).
\]

By Chebysev’s inequality, for every \( \epsilon' > 0 \),

\[
P \left( \left| \frac{1}{\tilde{T} - 2\tilde{H}} \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} \frac{R^2_{i,\tilde{t}+\tilde{H}}}{\mathbb{E}[R^2_{i,\tilde{t}+\tilde{H}}]} - 1 \right| \geq \epsilon' \right) \leq \frac{(\tilde{T}^{1/2} - \epsilon') \text{var} \left( \frac{1}{\tilde{T} - 2\tilde{H}} \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} \frac{R^2_{i,\tilde{t}+\tilde{H}}}{\mathbb{E}[R^2_{i,\tilde{t}+\tilde{H}}]} \right)}{\epsilon'^2} = O(\tilde{T}^{-1/2} - \epsilon')
\]

which converges to 0 as \( \tilde{T} \rightarrow \infty \).

\[\square\]

**Remark C.23.** Follow the proof of Lemma C.22, the same limit results apply to \( \tilde{R}^2_{\tilde{t},\tilde{t}+\tilde{H}} \) and \( \tilde{RV}^2_{i-\tilde{H},\tilde{t}} \).

**Lemma C.24.** Let

\[
S^2_R = \frac{1}{\tilde{T} - 2\tilde{H}} \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} (R_{i,\tilde{t}+\tilde{H}} - \overline{R}_{\tilde{H}})^2 \tag{C.80}
\]

\[
S^2_{RV} = \frac{1}{\tilde{T} - 2\tilde{H}} \sum_{i=\tilde{H}+1}^{\tilde{T} - \tilde{H}} (RV_{i-\tilde{H},\tilde{t}} - \overline{RV}_{\tilde{H}})^2 \tag{C.81}
\]

When \( d = 0 \), if \( 0 < \kappa < 1 \),

\[
\frac{S^2_R}{\text{var}(R_{i,\tilde{t}+\tilde{H}})} = 1 + o_p(\tilde{T}^{-1/2} + \epsilon'/2)
\]

\[
\frac{S^2_{RV}}{\text{var}(RV_{i\tilde{t},\tilde{t}})} = 1 + o_p(\tilde{T}^{-1/2} + \epsilon'/2)
\]

64
Proof. We can express
\[
\frac{S_{R}^{2}}{\text{var}(R_{i,i+\tilde{H}})} = \frac{1}{T-2\tilde{H}} \sum_{i=\tilde{H}+1}^{T-\tilde{H}} \frac{\text{var}(R_{i,i+\tilde{H}})}{\text{var}(R_{i,i+\tilde{H}})} \left( \frac{\text{cov}(R_{i,i+\tilde{H}})}{\text{var}(R_{i,i+\tilde{H}})} \right) - \left( \frac{\text{var}(R_{i,i+\tilde{H}})}{\text{var}(R_{i,i+\tilde{H}})} \right) \right) - \left( \frac{\text{cov}(R_{i,i+\tilde{H}})}{\text{var}(R_{i,i+\tilde{H}})} \right)
\]
From Lemma C.22 as well as Lemma C.19 if \(0 < \kappa < 1\) with \(\epsilon'\) chosen so that \(-1/2+\epsilon'/2 < 0\),
\[
\frac{S_{R}^{2}}{\text{var}(R_{i,i+\tilde{H}})} = (1 + o_{p}(\tilde{T}^{-1/2+\epsilon'/2})) \left( 1 + \left( \frac{\text{cov}(R_{i,i+\tilde{H}})}{\text{var}(R_{i,i+\tilde{H}})} \right) \right) - (1 + o_{p}(\tilde{T}^{-1/2+\epsilon'/2})) \left( \frac{\text{var}(R_{i,i+\tilde{H}})}{\text{var}(R_{i,i+\tilde{H}})} \right)
\]
The proof for \(\frac{S_{RV}^{2}}{\text{var}(R_{i,i+\tilde{H}})}\) is similar to \(S_{R}^{2}\) and thus is skipped.

\[
\text{Lemma C.25. If } d = 0,
\]
\[
\frac{S_{R}^{2}}{\tilde{H}} \xrightarrow{p} m \left( \text{var}(r_{0}) + 2 \sum_{k=1}^{h} \gamma_{r}(k) \right) \quad (C.82)
\]
\[
\frac{S_{RV}^{2}}{\tilde{H}} \xrightarrow{p} m \left( \text{var}(r_{0}^{2}) + 2 \sum_{k=1}^{h} \tilde{\gamma}_{r}(k) \right) \quad (C.83)
\]
where \(\gamma_{r}(k) = \text{cov}(r_{t}, r_{t+k})\) and \(\tilde{\gamma}_{r}(k) = \text{cov}(r_{t}^{2}, r_{t+k}^{2})\).

Proof. When \(d = 0\), \(r_{t}\) is \(h\)-dependent, hence [C.39] is equal to
\[
\text{var}(R_{i,i+\tilde{H}}) = \tilde{H} m \left[ \text{var}(r_{0}) + 2 \sum_{k=1}^{h} \left( 1 - \frac{k}{\tilde{H} m} \right) \gamma_{r}(k) \right]
\]
Therefore,
\[
\frac{S_{R}^{2}}{\text{var}(R_{i,i+\tilde{H}})} = \frac{S_{R}^{2}}{\tilde{H}} \cdot \frac{1}{m \left[ \text{var}(r_{0}) + 2 \sum_{k=1}^{h} (1 - \frac{k}{\tilde{H} m}) \gamma_{r}(k) \right]}
\]
As \(\tilde{H} \to \infty\), \(\sum_{k=1}^{h} \left( 1 - \frac{k}{\tilde{H} m} \right) \gamma_{r}(k) \to \sum_{k=1}^{h} \gamma_{r}(k)\). By Lemma C.24 for every \(\tilde{H} + h/m < \tilde{T} - 2\tilde{H}\) and every \(\epsilon' > 0\),
\[
\frac{S_{R}^{2}}{\tilde{H}} \cdot \frac{1}{m \left[ \text{var}(r_{0}) + 2 \sum_{k=1}^{h} \gamma_{r}(k) \right]} = 1 + o_{p}(\tilde{T}^{-1/2+\epsilon'/2})
\]
65
Therefore,
\[
\frac{S_R^2}{H} \xrightarrow{P} m \left( \text{var}(r_0) + 2 \sum_{k=1}^{h} \gamma_r(k) \right).
\]

The steps to obtain the limiting result for \(S_{RV}^2\) is similar to \(S_R^2\) and thus is skipped here. \(\square\)

**Remark C.26.** Follow Remark C.23 and the proofs of Lemmas C.24 and C.25, the same limiting results apply to \(S_{\tilde{R}}^2\), where \(S_{\tilde{R}}^2 = \frac{1}{T-2H} \sum_{i=H+1}^{\tilde{T}-\tilde{H}} \left( \tilde{R}_{i+\tilde{H}} - \tilde{R}_{\tilde{H}} \right)^2 \).

**Proof of Theorem 4.1**

**Proof.** Let
\[
\tilde{C}_T = \frac{1}{T-2H} \sum_{i=H+1}^{\tilde{T}-\tilde{H}} \left( \tilde{R}_{i+\tilde{H}} - \tilde{R}_{\tilde{H}} \right) \left( \tilde{R}_{\tilde{V}_{i+\tilde{H}}} - \tilde{R}_{\tilde{V}_{\tilde{H}}} \right),
\] (C.84)
and the theoretically-centered sample covariance as
\[
\tilde{C}_T^* = \frac{1}{T-2H} \sum_{i=H+1}^{\tilde{T}-\tilde{H}} \left( \tilde{R}_{i+\tilde{H}} - \mathbb{E}[\tilde{R}_{i+\tilde{H}}] \right) \left( \tilde{R}_{\tilde{V}_{i+\tilde{H}}} - \mathbb{E}[\tilde{R}_{\tilde{V}_{i+\tilde{H}}}] \right).
\]

Let
\[
S_R^2 = \frac{1}{T-2H} \sum_{i=H+1}^{\tilde{T}-\tilde{H}} \left( \tilde{R}_{i+\tilde{H}} - \tilde{R}_{\tilde{H}} \right)^2,
\]
and
\[
S_{RV}^2 = \frac{1}{T-2H} \sum_{i=H+1}^{\tilde{T}-\tilde{H}} \left( \tilde{R}_{\tilde{V}_{i+\tilde{H}}} - \tilde{R}_{\tilde{V}_{\tilde{H}}} \right)^2.
\]

We can then express
\[
\tilde{\rho}_{\tilde{H},\tilde{V}} = \frac{\tilde{C}_T}{S_RS_{RV}} = \frac{\tilde{C}_T - \tilde{C}_T^* + \tilde{C}_T^*}{S_RS_{RV}} = \frac{\tilde{C}_T^*}{S_RS_{RV}} + \frac{\tilde{C}_T - \tilde{C}_T^*}{S_RS_{RV}}.
\] (C.85)

By Lemma C.28 and Theorem C.29 for every \(\kappa \in (0, 1/3)\)
\[
\sqrt{\frac{T}{H^3}} \tilde{C}_T^* \xrightarrow{(d)} N \left( 0, \frac{2}{3} m^2 \sum_{|u| \leq h \atop |v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2) \right).
\] (C.86)
From Lemma C.25 and Remark C.26, when \(d = 0\), \(\frac{S_r}{H} \xrightarrow{p} m \left( \text{var}(r_0) + 2 \sum_{k=1}^{h} \gamma_r(k) \right)\) and \(\frac{S_R^2}{H} \xrightarrow{p} m \left( \text{var}(r_0^2) + 2 \sum_{k=1}^{h} \tilde{\gamma}_r(k) \right)\). Therefore, by Slutsky’s Theorem, it follows that for every \(\kappa \in (0, 1/3)\),

\[
\left( \frac{\sqrt{T}}{H^3} \right)^{\tilde{C}_T} = \sqrt{T^{1-\kappa}} \frac{\tilde{C}_T}{\sqrt{S_R^2 S_{RV}}} \xrightarrow{d} N \left( 0, \frac{S^2}{A_1 A_2} \right), \tag{C.87}
\]

where \(S^2 = \frac{2}{3} \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u)\text{cov}(r_v^2, r_0^2)\), \(A_1 = \text{var}(r_0) + 2 \sum_{k=1}^{h} \text{cov}(r_0, r_k)\), and \(A_2 = \text{var}(r_0^2) + 2 \sum_{k=1}^{h} \text{cov}(r_0^2, r_k^2)\).

Next, we show for \(0 < \kappa < 1/3\), \(\sqrt{T^{1-\kappa}} \left( \frac{\tilde{C}_T - \tilde{C}_T^*}{S_R S_{RV}} \right) \xrightarrow{} 0\).

We can express

\[
\tilde{C}_T - \tilde{C}_T^* = \frac{1}{T - 2H} \sum_{i=H+1}^{T-H} \left[ \left( \tilde{R}_{i,i+H} - \frac{\tilde{R}_{i,i+H}}{\tilde{R}_{i,i+H}} \right) \left( \tilde{R}_{i,i+H} - \frac{\tilde{R}_{i,i+H}}{\tilde{R}_{i,i+H}} \right) - \left( \tilde{R}_{i,i+H} - \mathbb{E}[\tilde{R}_{i,i+H}] \right) \left( \mathbb{E}[\tilde{R}_{i,i+H}] - \tilde{R}_{i,i+H} \right) \right]
\]

From Lemma C.19 and Remark C.21, when \(d = 0\), as \(\tilde{T} \to \infty\), \(\forall 0 < \kappa < 1/3\),

\[
\tilde{R}_{i,i+H} = \mathbb{E}[\tilde{R}_{i,i+H}] + \mathbb{E}[\tilde{R}_{i,i+H}] o_p(\tilde{T}^{-1/2+\epsilon'/2}) = \mathbb{E}[\tilde{R}_{i,i+H}] + o_p(\tilde{T}^{-1/2+\epsilon'/2})
\]

\[
\tilde{R}_{i,i+H} = \mathbb{E}[\tilde{R}_{i,i+H}] + \mathbb{E}[\tilde{R}_{i,i+H}] o_p(\tilde{T}^{-1/2+\epsilon'/2}) = \mathbb{E}[\tilde{R}_{i,i+H}] + o_p(\tilde{T}^{-1/2+\epsilon'/2})
\]

It follows

\[
\tilde{C}_T - \tilde{C}_T^* = \left( \mathbb{E}[\tilde{R}_{i,i+H}] - \tilde{R}_{i,i+H} \right) \left( \mathbb{E}[\tilde{R}_{i,i+H}] - \tilde{R}_{i,i+H} \right) = o_p(\tilde{T}^{2\kappa-1+\epsilon'}). \tag{C.88}
\]

Since

\[
\sqrt{T^{1-\kappa}} \left( \frac{\tilde{C}_T - \tilde{C}_T^*}{S_R S_{RV}} \right) = \sqrt{T^{1-\kappa}} \left( \frac{\tilde{C}_T - \tilde{C}_T^*}{S_R S_{RV}} \right) \left( \frac{1}{H} \right), \tag{C.89}
\]

67
Lemma C.27. Let \( \tilde{R}_k^* \) and \( RV_k^* \) be the centered versions of \( R_k \) and \( RV_k \): \( \tilde{R}_k^* = R_k - \mathbb{E}[R_k] \), \( RV_k^* = RV_k^* - \mathbb{E}[RV_k] \), where \( \tilde{R}_k = \sum_{j=km+h+1}^{(k+\tilde{H})m} r_j \) and \( RV_k = \sum_{j=km+1}^{(k+\tilde{H})m} r_j^2 \). Let \( r_j^* = r_i - \mathbb{E}[r_i] \) and \( s_j^* = r_j^2 - \mathbb{E}[r_j^2] \).

(i) if \( k > \tilde{H} \), then

\[ \text{cov}(\tilde{R}_0^* RV_0^*, \tilde{R}_k^* RV_k^*) = 0 \]  \hspace{1cm} (C.91)

(ii) if \( 0 \leq k \leq \tilde{H} \), then

\[ \text{cov}(\tilde{R}_0^* RV_0^*, \tilde{R}_k^* RV_k^*) = c_k + O(1) \]  \hspace{1cm} (C.92)

where \( c_k \) is given by \( (C.105) \).

Proof. Applying Proposition 3.2.1 in Peccati and Taqqu (2011) and properties of cumulants, we show that for \( k \geq 0 \),

\[
\begin{align*}
\text{cov}(\tilde{R}_0^* RV_0^*, \tilde{R}_k^* RV_k^*) &= \sum_{i=0}^{\tilde{R}_m} \sum_{j=-\tilde{H}_{m+1}}^{(k+\tilde{H})m} \sum_{u=km+h+1}^{(k+H)m+1} \sum_{v=(k-H)m+1}^{k} \text{cov}(r_i^*, s_j^*, r_u^* s_v^*) \\
&= \sum_{i=0}^{\tilde{R}_m} \sum_{j=-\tilde{H}_{m+1}}^{(k+\tilde{H})m} \sum_{u=km+h+1}^{(k+H)m+1} \sum_{v=(k-H)m+1}^{k} \text{cum}(r_i^*, r_j^2, r_u^2, r_v^2) \quad \text{(cum)} \\
&+ \sum_{i=0}^{\tilde{R}_m} \sum_{j=-\tilde{H}_{m+1}}^{(k+\tilde{H})m} \sum_{u=km+h+1}^{(k+H)m+1} \sum_{v=(k-H)m+1}^{k} \text{cov}(r_i^*, r_u) \text{cov}(r_j^2, r_v^2) \quad \text{(cov1)} \\
&+ \sum_{i=0}^{\tilde{R}_m} \sum_{j=-\tilde{H}_{m+1}}^{(k+\tilde{H})m} \sum_{u=km+h+1}^{(k+H)m+1} \sum_{v=(k-H)m+1}^{k} \text{cov}(r_u, r_j^2) \text{cov}(r_i, r_v^2) \quad \text{(cov2)}
\end{align*}
\]
Consider the term \((\text{cum})\).

Since \(|i - j| > h\), \(|u - j| > h\), and \(|u - v| > h\), \(r_i\) is independent of \(r_j\) and \(r_u\) is independent of \(r_j\) and \(r_v\).

A. If \(k > \tilde{H}\), then \(r_i\) is independent of \(r_u\). Also \(r_j\) and \(r_v\) are independent of \(r_u\), hence the elements of \(\{r_i, r_j, r_u\}\) are independent of \(r_u\), and thus \(\text{cum}(r_i, r_j^2, r_u, r_v^2) = 0\). Therefore, if \(k > \tilde{H}\), the term \((\text{cum})\) is zero.

B. If \(0 \leq k \leq \tilde{H}\), then \((k - \tilde{H})m + 1 \leq 1\). We can split the range of \(v\) into two parts:

(i) if \((k - \tilde{H})m + 1 \leq v \leq 0\), then \(r_i\) is also independent of \(r_v\), and the elements of \(\{r_i, r_u\}\) are independent of the elements \(\{r_j, r_v\}\). Hence \(\text{cum}(r_i, r_j^2, r_u, r_v^2) = 0\).

(ii) if \(1 \leq v \leq km\), the \((\text{cum})\) term equals

\[
\sum_{i=h+1}^{\tilde{H}m} \sum_{j=-Hm+1}^{0} \sum_{u=km+h+1}^{(k+\tilde{H})m} \sum_{v=1}^{km} \text{cum}(r_i, r_j^2, r_u, r_v^2) . \tag{C.93}
\]

There are terms in \((C.93)\) with \(\text{cum}(r_i, r_j^2, r_u, r_v^2) = 0\).

If \(-\tilde{H}m + 1 \leq j \leq -h\) or \(v \geq h + 1\), then \(r_j\) is independent of \(r_v\), and the elements of \(\{r_i, r_u, r_v\}\) are independent of \(\{r_j\}\). Thus \(\text{cum}(r_i, r_j^2, r_u, r_v^2) = 0\). Therefore, after excluding those terms with \(\text{cum}(r_i, r_j^2, r_u, r_v^2) = 0\), \((C.93)\) becomes

\[
\sum_{i=h+1}^{\tilde{H}m} \sum_{j=-h+1}^{0} \sum_{u=km+h+1}^{(k+\tilde{H})m} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v^2) . \tag{C.94}
\]

Next we check if the elements of \(\{r_i, r_u\}\) are independent of the elements of \(\{r_j, r_v\}\) in \((C.94)\).

If \(u > 2h\), then \(r_u\) is independent of \(r_v\) and \(r_j\). Also if \(i > 2h\), then \(r_i\) is independent of \(r_j\) and \(r_v\). It follows \(\{r_i, r_u\}\) is independent of \(\{r_j, r_v\}\) and \(\text{cum}(r_i, r_j^2, r_u, r_v^2) = 0\). Hence we have

\[
\sum_{i=2h+1}^{\tilde{H}m} \sum_{j=-h+1}^{0} \sum_{u=(km+h+1) v=1}^{(k+\tilde{H})m} \sum_{v=1}^{2h+1} \text{cum}(r_i, r_j^2, r_u, r_v^2) = 0 . \tag{C.95}
\]

In the following, we discuss two cases regarding the relationship between \(km+h+1\) and \(2h+1\):
If \((km + h + 1) \lor (2h + 1) = 2h + 1\), we can express \((C.94)\) as
\[
\tilde{H}_m \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=km+h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) \quad (C.96)
\]
\[
+ \tilde{H}_m \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=2h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) \quad (C.97)
\]

After excluding the terms in \((C.95)\) from \((C.97)\), \((C.94)\) becomes
\[
\tilde{H}_m \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=km+h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) \quad (C.98)
\]
\[
+ 2h \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=2h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) \quad (C.99)
\]

In \((C.98)\), if \(i > 3h\) then \(r_i\) is independent of \(r_u\) and the elements of \(\{r_i, r_j\}\) are independent of the elements of \(\{r_u, r_v\}\). Also in \((C.99)\), if \(u > 3h\), then \(r_i\) is independent of \(r_u\) and the set \(\{r_i, r_j, r_v\}\) is independent of \(r_u\). After excluding these terms with \(\text{cum}(r_i, r_j^2, r_u, r_v) = 0\), sum of \((C.98)\) and \((C.99)\) become
\[
3h \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=km+h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) \\
+ 2h \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=2h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) = O(1) \quad (C.100)
\]
uniformly w.r.t. \(k\) and \(\tilde{H}\).

Similarly, if \((km + h + 1) \lor (2h + 1) = km + h + 1\). Hence we can express \((C.95)\) as
\[
\tilde{H}_m \sum_{i=2h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=km+h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) \quad (C.101)
\]

After excluding the terms in \((C.101)\) from \((C.94)\), \((C.94)\) becomes
\[
2h \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{h} \sum_{u=km+h+1}^{v=1} \sum_{v=1}^{h} \text{cum}(r_i, r_j^2, r_u, r_v) \quad (C.102)
\]
If \( u > 3h \), then \( r_u \) is independent of \( r_i \) and \( r_v \), and the elements of \( \{r_i, r_j, r_v\} \) are independent of \( r_u \). Therefore, (C.102) becomes

\[
2h \sum_{i=h+1}^{2h} \sum_{j=-h+1}^{0} \sum_{u=km+h+1}^{(k+H)m} \sum_{v=1}^{h} \text{cum}(r_i, r_j, r_u, r_v) = O(1) ,
\]

(C.103)

From the results of (C.100) and (C.103), it follows \( \text{cum} = O(1) \), uniformly w.r.t. \( k \) and \( \tilde{H} \).

Next, we consider the sum \( \text{cov}2 \). Since \( u > km + h + 1 \) and \( j \leq 0 \), \( r_u \) is independent of \( r_j \), and thus \( \text{cov}(r_u, r_j^2) = 0 \). It follows

\[
\text{cov2} = \tilde{H}m \sum_{i=h+1}^{\tilde{H}m} \sum_{v=(k-\tilde{H})m+1}^{km} \text{cov}(r_i, r_v^2) \sum_{j=-\tilde{H}m+1}^{0} \sum_{u=km+h+1}^{(k+\tilde{H})m} \text{cov}(r_u, r_j^2) = 0 .
\]

(C.104)

Finally, consider the sum \( \text{cov}1 \).

If \( k > \tilde{H} \), then \( r_i \) and \( r_u \) are independent and thus \( \text{cov}(r_i, r_u) = 0 \). It follows that \( \text{cov1} = 0 \). Also \( \text{cov}2 \) and \( \text{cum} \) are both zero if \( k > \tilde{H} \). Therefore, if \( k > \tilde{H} \), \( \text{cov}(\tilde{R}_0^*RV_0^*, \tilde{R}_k^*RV_k^*) = 0 \).

If \( 0 \leq k \leq \tilde{H} \), since \( \{r_i\} \) are \( h \)-dependent, we can express

\[
\text{cov1} = \tilde{H}m \sum_{i=h+1}^{\tilde{H}m} \sum_{j=-\tilde{H}m+1}^{0} \sum_{u=km+h+1}^{(k+\tilde{H})m} \sum_{v=(k-\tilde{H})m+1}^{km} \text{cov}(r_i, r_u)\text{cov}(r_j^2, r_v^2)
\]

\[= \tilde{H}m \sum_{u=(k-\tilde{H})m+h+1}^{(k+\tilde{H})m-1+h} \sum_{i=h+1}^{(k+\tilde{H})m} 1\{km + h + 1 \leq i + u \leq (k + \tilde{H})m\} \text{cov}(r_0, r_u)\]

\[\times \sum_{v=(k-\tilde{H})m+1}^{(k+\tilde{H})m-1+h} \sum_{j=-\tilde{H}m+1}^{0} 1\{(k - \tilde{H})m + 1 \leq j + v \leq km\} \text{cov}(r_0^2, r_v^2)\]

\[= \sum_{u=(k-\tilde{H})m+h+1}^{(k+\tilde{H})m-1+h} \sum_{j=-\tilde{H}m+1}^{(k+\tilde{H})m} (\tilde{H}m - h - |km - u|)^+ \text{cov}(r_0, r_u)\]

\[\times \sum_{v=(k-\tilde{H})m+1}^{(k+\tilde{H})m-1+h} (\tilde{H}m - |km - v|)^+ \text{cov}(r_0^2, r_v^2) .
\]

(C.105)
Let $c_k = \text{cov}$. By the results below (C.103) as well as (C.104), it follows that 
\[
\text{cov}(\tilde{R}_0^* R V_0^*, \tilde{R}_k^* R V_k^*) = c_k + O(1).
\]

**Lemma C.28.** Define $\tilde{C}_T^*, V_T, W_T$ and $\tilde{\rho}_{T, T}$ by

\[
V_T = \tilde{T}^{-1} \sum_{k=1}^{\tilde{T}} (\tilde{R}_k^*)^2, \quad W_T = \tilde{T}^{-1} \sum_{k=1}^{\tilde{T}} (R V_k^*)^2, \quad \tilde{C}_T^* = \tilde{T}^{-1} \sum_{k=1}^{\tilde{T}} \tilde{R}_k^* R V_k^*, \quad \tilde{\rho}_{T, T} = \frac{\tilde{C}_T^*}{\sqrt{V_T W_T}}.
\]

\[
\lim_{H \to \infty} \tilde{T} H^{-3} \text{var}(\tilde{C}_T^*) = \frac{2}{3} \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2).
\]

**Proof.** By definition,

\[
\text{var}(\tilde{C}_T^*) = \frac{1}{T} \text{var} \left( \tilde{R}_0^* R V_0^* \right) + \frac{2}{T^2} \sum_{k=1}^{\tilde{T}} (\tilde{T} - k) \text{cov} \left( \tilde{R}_0^* R V_0^*, \tilde{R}_k^* R V_k^* \right)
\]

From Lemma C.27, if $k > H$, \( \text{cov} \left( \tilde{R}_0^* R V_0^*, \tilde{R}_k^* R V_k^* \right) = 0 \). Hence we can express

\[
T \tilde{T} H^{-3} \text{var}(\tilde{C}_T^*) = \tilde{T} H^{-3} \text{var}(\tilde{R}_0^* R V_0^*) + 2 \tilde{T} H^{-3} \sum_{k=1}^{\tilde{H}} \left( 1 - \frac{k}{\tilde{T}} \right) (c_k + O(1))
\]

\[
= \tilde{T} H^{-3} \text{var}(\tilde{R}_0^* R V_0^*) + 2 \tilde{T} H^{-3} \sum_{k=1}^{\tilde{H}} c_k - 2 \tilde{T} H^{-3} \sum_{k=1}^{\tilde{T}} \frac{k}{\tilde{T}} c_k
\]

\[
+ 2 \tilde{T} H^{-3} \sum_{k=1}^{\tilde{H}} \left( 1 - \frac{k}{\tilde{T}} \right) O(1)
\]

From (C.105) as well as the results below (C.103) and (C.104), if $k = 0$,

\[
\lim_{H \to \infty} \text{var}(\tilde{R}_0^* R V_0^*) = O(\tilde{H}^2)
\]

Thus

\[
\tilde{T} H^{-3} \text{var}(\tilde{R}_0^* R V_0^*) = O(\tilde{H}^{-1})
\]

72
If $1 \leq k \leq \tilde{H}$, then

$$|c_k| = \tilde{H}^2 m^2 \left| \sum_{|u| \leq h} \left(1 - \frac{h}{\tilde{H} m} - \frac{\left|km - u\right|}{\tilde{H} m} \right)^+ \text{cov}(r_0, r_u) \times \sum_{|v| \leq h} \left(1 - \frac{|km - v|}{\tilde{H} m} \right)^+ \text{cov}(r_0^2, r_v^2) \right|$$

$$\leq \tilde{H}^2 m^2 \sum_{|u| \leq h} \sum_{|v| \leq h} \left| \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2) \right|$$

(C.109)

Therefore,

$$\tilde{H}^{-3} \left| \sum_{k=1}^{\tilde{H}} \frac{k}{T} c_k \right| = \frac{\tilde{H}^{-2}}{T} \left| \sum_{k=1}^{\tilde{H}} \frac{k}{T} c_k \right| \leq \frac{\tilde{H}^{-2}}{T} \sum_{k=1}^{\tilde{H}} |c_k| \leq \frac{\tilde{H}^{-2}}{T} \tilde{H}O(\tilde{H}^2) = O(\tilde{H}/\tilde{T})$$.

(C.110)

Also

$$\tilde{H}^{-3} \sum_{k=1}^{\tilde{H}} \left(1 - \frac{k}{\tilde{T}}\right) = \frac{1}{H^2} - \frac{1}{2TH^2} - \frac{1}{2TH} \to 0$$,

(C.111)

as $\tilde{H} \to \infty$. Therefore, by the results of (C.108), (C.110) and (C.111), Equation (C.107) can be represented as

$$\tilde{T} \tilde{H}^{-3} \text{var}(\hat{C}_T) = 2\tilde{H}^{-3} \sum_{k=1}^{\tilde{H}} c_k + O(\tilde{H}^{-1} + \tilde{H}/\tilde{T})$$.

(C.112)

Consider the term

$$\tilde{H}^{-3} \sum_{k=1}^{\tilde{H}} c_k = \tilde{H}^{-3} \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2) \sum_{|u| \leq h} \sum_{|v| \leq h} \left(\tilde{H} m - |km - u|\right)^+ \left(\tilde{H} m - |km - v|\right)^+$$

Then

$$\lim_{\tilde{H} \to \infty} \tilde{H}^{-3} \sum_{k=1}^{\tilde{H}} c_k = \lim_{\tilde{H} \to \infty} \sum_{k=1}^{\tilde{H}} \sum_{|u| \leq h} \sum_{|v| \leq h} \left(1 - \frac{|k - u|}{\tilde{H}}\right)^+ \left(1 - \frac{|k - v|}{\tilde{H}}\right)^+ \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2)$$

$$= m^2 \lim_{\tilde{H} \to \infty} \sum_{k=1}^{\tilde{H}} \left(1 - \frac{k}{\tilde{H}}\right)^2 \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2)$$

$$\to m^2 \int_0^1 (1 - t)^2 dt \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2)$$

$$= \frac{1}{3} m^2 \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2)$$.

73
It follows
\[
\lim_{\tilde{H} \to \infty} \tilde{T} \tilde{H}^{-3} \text{var}(\tilde{C}_T^*) = \frac{2}{3} m^2 \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2).
\]

\[\square\]

**Theorem C.29.** Let \( T_k = \tilde{R}_k^* \text{RV}_k^* = \sum_{i=km+1}^{(k+\tilde{H})m} (r_i - E[r_0]) \sum_{j=(k-\tilde{H})m+1}^{km} (r_j^2 - E[r_0^2]) \), where \( \tilde{R}_k^* \) and \( \text{RV}_k^* \) are defined in Lemma C.27 and \( 1 \leq k \leq \tilde{T} \). If \( d = 0 \) and \( \tilde{H}^3/\tilde{T} \to 0 \), then
\[
(\tilde{T} \tilde{H}^3)^{-1/2} \sum_{k=1}^{\tilde{T}} T_k \xrightarrow{d} N \left( 0, \frac{2}{3} m^2 \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2) \right).
\]

**Proof.** In order to prove the central limit theorem, we apply the theorem of [Berk(1973)]. Let \( \text{cst} \) be a generic positive constant. We must check the following conditions

(i) There exists an integer \( p > 2 \) such that for \( 1 \leq k \leq \tilde{T} \), \( E[|T_k|^p] \leq \text{cst} \tilde{H}^p \) and \( \tilde{H}^{(2p-2)/(p-2)}/\tilde{T} \to 0 \).

(ii) \( \text{var}(T_1 + \cdots + T_k) \leq \text{cst} \tilde{H}^3 \), for \( 1 \leq k \leq \tilde{T} \).

(iii) \( \lim_{\tilde{T} \to \infty} (\tilde{T} \tilde{H}^3)^{-1} \text{var}(T_1 + \cdots + T_{\tilde{T}}) = \text{cst} \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2) \).

To check conditions (ii) and (iii), first we derive the expression for
\[
\text{var} \left( \sum_{j=1}^{k} T_j \right) = k \text{var}(T_0) + 2 \sum_{j=1}^{k} (k - j) \text{cov}(T_0, T_j).
\]

By Lemma C.27 if \( k > \tilde{H} \), \( \text{cov}(T_0, T_k) = 0 \). If \( k = 0 \), \( \text{var}(T_0) = O(\tilde{H}^2) \). For \( 1 \leq k \leq \tilde{H} \),
\[
\text{cov}(T_0, T_k) = c_k + O(1). \text{ Thus by (C.109),}
\]

\[
\sum_{j=1}^{k} (k - j) \text{cov}(T_0, T_j) = \sum_{j=1}^{k \wedge \bar{H}} (k - j) (c_j + O(1))
\]

\[
= k \sum_{j=1}^{k \wedge \bar{H}} \left(1 - \frac{j}{k}\right) c_k + O(k \wedge \bar{H})
\]

\[
\leq k \sum_{j=1}^{k \wedge \bar{H}} |c_j| + O(k \wedge \bar{H})
\]

\[
\leq k O(\bar{H}^3). 
\]

It follows that

\[
\text{var} \left( \sum_{j=1}^{k} T_j \right) \leq k O(\bar{H}^3), \tag{C.113}
\]

which proves condition (iii).

Let \( \tilde{C}_T^* = \frac{1}{T} \sum_{k=1}^{T} T_k \). Since \( \text{var}(T_1 + T_2 + \cdots + T_T) = \bar{T}^2 \text{var}(\tilde{C}_T^*) \). It follows from Lemma C.28

\[
(\bar{T}K^3)^{-1} \text{var} \left( \sum_{k=1}^{T} T_k \right) \rightarrow \frac{2}{3} \sum_{|u| \leq h} \sum_{|v| \leq h} \text{cov}(r_0, r_u) \text{cov}(r_0^2, r_v^2). 
\]

This proves condition (iii).

Next, we check condition (i). Since \( T_0 \) has zero mean, we can express

\[
\mathbb{E}[T_0^4] = \text{cum}(T_0, T_0, T_0, T_0) + 3 \text{var}(T_0).
\]

Thus it suffices to prove that \( \text{cum}(T_0, T_0, T_0, T_0) = O(K^4) \). This will show that (i) holds with \( p = 4 \) and the condition on \( K \) is \( K^3/n \rightarrow 0 \). Applying the properties of cumulants, we have

\[
\text{cum}(T_0, T_0, T_0, T_0) = \sum_{i=h+1}^{K} \sum_{j=-K+1}^{0} \sum_{k=h+1}^{K} \sum_{\ell=-K+1}^{0} \sum_{u=h+1}^{K} \sum_{v=-K+1}^{0} \sum_{w=h+1}^{K} \sum_{z=-K+1}^{0} \text{cum}(\bar{r}_i s_j^2, \bar{r}_k s_\ell^2, \bar{r}_u s_v^2, \bar{r}_w s_z^2)
\]
By Proposition 3.2.1 by Pecatti and Taqqu (2011) and that \( r_i, r_k, r_u, \) and \( r_w \) are independent of \( r_j, r_\ell, r_v, \) and \( r_z, \) we obtain that for \( i > h, k > h, u > h, \) and \( w > h, \)

\[
\text{cum}(T_0, T_0, T_0, T_0) = (K - h)^2 K^2 \text{var}^2(r_0) \text{var}^2(r_0^2) + (K - h) K \text{cum}(r_0, r_0, r_0, r_0) \text{cum}(r_0^2, r_0^2, r_0^2, r_0^2) \\
+ (K - h)^2 K \text{var}^2(r_0) \text{cum}(r_0^2, r_0^2, r_0^2, r_0^2) + (K - h) K^2 \text{var}^2(r_0^2) \text{cum}(r_0, r_0, r_0, r_0). 
\]

(C.114)

By Lemma C.30, \( \text{cum}(T_0, T_0, T_0, T_0) = O(K^4). \)

Lemma C.30. For every positive integer \( p, \) \( \mathbb{E}[r_i^p] < \infty. \)

Proof. From the return model, we can express the \( p^{th} \) moment of the return as

\[
\mathbb{E}[r_i^p] = \mathbb{E} \left[ \mu \Delta N(t) + \sum_{k=N(t-1)+1}^{N(t)} e_k \right]^p 
\]

The term with the highest moment on the RHS of the above equation is \( \mathbb{E}[((\Delta N(t))^p)], \) which can be represented as \( \mathbb{E}[\mathbb{E}[(\Delta N(t))^p|\Lambda]]. \) Conditional on \( \Lambda, \) \( \Delta N(t) \) is a Poisson random variable with mean \( \mathbb{E}[\Delta N(t)|\Lambda] = \lambda \int_{t-1}^{t} e^{Z_H(s)} ds. \) Thus \( \mathbb{E}[(\Delta N(t))^p|\Lambda] \) is the \( p^{th} \) moment of the Poisson variable \( \Delta N(t)|\Lambda. \)

Let \( \nu = \mathbb{E}[(\Delta N(t)|\Lambda] \) be the mean of \( \Delta N(t)|\Lambda. \) Since the non-centered \( p^{th} \) moment of a Poisson distribution is a polynomial in its mean with the highest order \( p, \) we can express \( \mathbb{E}[(\Delta N(t))^p|\Lambda] \) as the sum of the expected values of \( \nu \) with different exponentials. Hence \( \mathbb{E}[(\Delta N(t))^p] \) involves evaluating the term \( \mathbb{E}[^p]\nu]. \) Since \( Z_H \) is Gaussian stationary process, it has finite moments. It follows that \( \mathbb{E}[\nu^p] \) is finite and thus \( \mathbb{E}[(\Delta N(t))^p] \) is finite.

References

[1] T. Andersen, T. Bollerslev, F. X. Diebold, P. Labys, 2001. The Distribution of Realized Exchange Rate Volatility. Journal of the American Statistical Association, 96, 42-55.

[2] Y. Amihud, M. Hurvich, 2004. Predictive Regressions: A Reduced-Bias Estimation Method. Journal of Financial and Quantitative Analysis, 39(4), December, 813-841.
[3] R. Bansal, A. Yaron, 2004. Risks for the Long Run: A Potential Resolution of Asset Pricing Puzzles. *Journal of Finance*, **59**, 1481-1509.

[4] F. M. Bandi, B. Perron, 2006. Long Memory and the Relation Between Implied and Realized Volatility. *Journal of Financial Econometrics*, **4**(4), 636-670.

[5] F. M. Bandi, B. Perron, 2008. Long-run risk-return trade-offs. *Journal of Econometrics*, **143**, 349–374.

[6] T. Bollerslev, G. Tauchen, H. Zhou, 2009. Expected Stock Returns and Variance Risk Premia. *The Review of Financial Studies*, **22**(11), 4463-4492.

[7] T. Bollerslev, N. Sizova, G. Tauchen, 2012. Volatility in Equilibrium: Asymmetries and Dynamic Dependencies. *Review of Finance*, **16**, 31-80.

[8] J. Boudoukh, M. Richardson, 1993. Stock returns and inflation: a long-horizon perspective. *American Economic Review*, **83**(5), 1346–1355.

[9] J. Boudoukh, M. Richardson, R.F. Whitelaw, 2008. The Myth of Long-Horizon Predictability. *Review of Financial Studies*, **21**(4), 1577-1605.

[Berk(1973)] Kenneth N. Berk, 1973. A central limit theorem for \( m \)-dependent random variables with unbounded \( m \). *The Annals of Probability*, 1:352–354, 1973.

[10] J. Y. Campbell, R. Shiller, 1987. Cointegration and tests of present value models. *Journal of Political Economy*, **95**, 1062-1088.

[11] J. Y. Campbell, R. Shiller, 1988. The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies*, **1**, 195–228.

[12] J. Y. Campbell, M. Yogo, 2006. Efficient tests of stock return predictability. *Journal of Financial Economics*, **81**, 27-60.

[13] W. Cao, R. Deo, M. Hurvich, 2017. Drift in Transaction-Level Asset Price Models. *Journal of Time Series Analysis*, **38**, 769-790.

[14] W. Chen, R. Deo, Yanping Yi, 2013. Uniform Inference in Predictive Regression Models. *Journal of Business & Economic Statistics*, **31**(4), 525-533.
[15] F. Comte, E. Renault, 1998. Long Memory in Continuous-Time Stochastic Volatility Models. *Mathematical Finance, 8*(4), 291–323.

[16] R. Davidson, J.G. MacKinnon, 1999. The Size Distortion of Bootstrap Tests. *Econometric Theory, 15*, 361-376.

[17] R. Davidson, J.G. MacKinnon, 2006. The Power of Bootstrap and Asymptotic Tests. *Journal of Econometrics, 133*, 421-441.

[18] R.B. Davis, D.S. Harte, (1987). Test for Hurst effect. *Biometrika, 74*, 95-102.

[19] R. Deo, M. Hsieh, C. Hurvich, 2010. Long memory in intertrade durations, counts and realized volatility of NYSE stocks. *Journal of Statistical Planning and Inference, 140*(12), 3715-3733.

[20] R. Deo, C. Hurvich, P. Soulier, Y. Wang, 2009. Conditions for the propagation of memory parameter from durations to counts and realized volatility. *Econometric Theory, 25*(3), 764–792.

[21] I. Drechsler, A. Yaron, 2011. What’s Vol Got to Do with It. *The Review of Financial Studies, 24*(1), 1-45.

[22] R. F. Engle, J. R. Russell, 1998. Autoregressive conditional duration: a new model for irregularly spaced transaction data. *Econometrica, 66*(5), 1127-1162.

[23] E. Fama, K. French, 1988. Dividend yields and expected stock returns. *Journal of Financial Economics, 22*, 3–25.

[24] A. Kostakis, T. Magdalinosy, M. P. Stamatogiannis, 2018. Taking Stock of Long-Horizon Predictability Tests: Are Factor Returns Predictable ? Working paper.

[25] F. Mishkin, 1990. What does the term structure of interest rates tell us about future inflation ? *Journal of Monetary Economics, 70*, 1064–1072.

[26] F. Mishkin, 1992. Is the Fisher effect for real? *Journal of Monetary Economics, 30*(2), 195-215.

[27] V. Pipiras, M. Taqqu, 2017. Long-Range Dependence and Self-Similarity. *Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press.*
[28] G. Peccati and Murad S. Taqqu, 2011. *Wiener chaos: moments, cumulants and diagrams*, volume 1 of *Bocconi & Springer Series*. Springer, Milan; Bocconi University Press, Milan. A survey with computer implementation, Supplementary material available online.

[29] N. Sizova, 2013. Long-Horizon Return Regressions With Historical Volatility and Other Long-Memory Variables. *Journal of Business and Economic Statistics*, 31(4), 546-559.

[30] R. Stambaugh, 1999. Predictive regressions. *Journal of Financial Economics*, 54, 375-421.

[31] R. Valkanov, 2003. Long-horizon regressions: theoretical results and applications. *Journal of Financial Economics*, 68, 201-232.