



# **Temporal Aggregation of Random Walk Processes and Implications for Asset Prices**

**By**

**Yamin Ahmad  
(UW-Whitewater)**

**and**

**Ivan Paya  
(Lancaster University Management School)**

**Working Paper 14 - 01**

**University of Wisconsin – Whitewater**  
Department of Economics  
4<sup>th</sup> Floor Hyland Hall  
800 W. Main Street  
Whitewater, WI 53190

**Tel: (262) 472 -1361**

# Temporal Aggregation of Random Walk Processes and Implications for Asset Prices

Yamin S Ahmad\*

University of Wisconsin - Whitewater

Ivan Paya†

Lancaster University Management School, Economics Department

## Abstract

This paper examines the impact of time averaging and interval sampling data assuming that the data generating process for a given series follows a random walk with uncorrelated increments. We provide expressions for the corresponding variances, and covariances, for both the levels and differences of the aggregated series, demonstrating how the degree of temporal aggregation impacts these particular properties. Moreover, we analytically derive any differences that arise between the aggregated series and its disaggregated counterpart, and show that they can be decomposed into a distortionary and small sample effect. We also provide exact expressions for the variance and sharpe ratios, and correlation coefficients for any level of aggregation. We discuss our results in the context of asset prices, which have utilized these extensively.

*JEL Classification:* F47 C15 C32

*Keywords:* Temporal Aggregation, Random walk, variance ratio, sharpe ratio

Current version: March 7th, 2014

---

\* Department of Economics, 800 W Main Street, Whitewater, WI 53190 Email: ahmady@uww.edu, Homepage: <http://facstaff.uww.edu/ahmady/> Tel: (262) 472 5576, Fax: (262) 472 4683

† Lancaster University Management School, Lancaster, LA1 4YX, UK Email: i.paya@lancaster.ac.uk

# 1 Introduction

The problems of temporal aggregation of time series models and the biases that result from utilizing this type of procedure have been known since at least the 1960's. Yet they have had relatively little impact on the econometric literature and the approach adopted by applied economists attempting to address questions of interest. Macroeconomists and policy makers, for example, often model the dynamics of key macroeconomic variables at a quarterly frequency, without citing any reasons governing that choice, *a priori*, aside from the availability of data. This is of course understandable, given the limited resources that can be allocated towards the collection of data by the various institutions that do compile the data used by academics and policy makers alike. However, in many circumstances, higher frequency data is available and simply modeling the data generating processes of key macroeconomic and financial variables using lower frequency data, which have been temporally aggregated, fails to address two important questions. The first is the appropriateness of the chosen frequency in characterizing the 'true' data generating process. The second is whether a particular 'natural' interval exists, whereby if one were to collect data at that interval, it would approximate the natural frequency driving the particular process (see Brewer, 1973). Any progress in attempting to answer these types of questions relies on understanding the implications of temporally aggregating data. In particular, an essential ingredient is to understand how the degree of temporal aggregation and sample size relates to the properties of a derived time series process, as this then has implications for estimation and inference. Our contribution in this paper goes to the heart of this issue, as we demonstrate exactly how the degree of temporal aggregation affects the variance, covariance and autocorrelation of an aggregated time series, and we concomitantly examine the implications for variance ratio tests and Sharpe ratios that have been used in the macroeconomic and financial literatures.

One of the main contributions of temporal aggregation in the economics literature can be attributed

to Holbrook Working (1960), who examined biases that arose as a result of time averaging data. Working's (1960) study focused on the impact of serial correlation in autoregressions containing price differentials, where he demonstrated and derived analytical expressions for a bias that emerges through temporal aggregation for the variance and covariance of a first differenced price series. Subsequent research on the theoretical front has focused on deriving the implications of temporal aggregation as they pertain to the appropriate way to characterize the data generating process for an aggregated series.<sup>1</sup> Even up until a decade ago, theoretical work on temporal aggregation has derived the implications for a number of properties of an aggregated series, like persistence and half-life (Taylor, 2001).

In most empirical studies, the econometrician knows relatively little about the properties of an underlying data generating process, but uses available data to estimate and infer something about those properties. The issue of temporal aggregation has been taken into account when analyzing the persistence of macroeconomic series for policy analysis (Rossana and Seater, 1992, 1995; Paya et al., 2007) or when testing specific models that otherwise would suggest rejections of theories such as the Permanent Income Hypothesis (Haug, 1991, Christiano et al., 1991), or Purchasing Power Parity (Choi et al., 2006; Paya and Peel, 2006, Ahmad and Craighead, 2011).<sup>2</sup> A general conclusion of the studies cited above is that temporal aggregation tends to induce additional persistence in the series. However, as insightful as all these results are, the question of how temporal aggregation affects a time series process for a specific level of aggregation and sample size still remains an open one.

Consequently in this paper, we extend the line of research initiated by Working (1960) by examining

---

<sup>1</sup>For example, Amemiya and Wu (1972) show how temporal aggregation induces moving average terms in a purely autoregressive model. Brewer (1973) computes the number of terms to which more general  $ARMA(p, q)$  processes approach as the level of aggregation increases. Tiao (1972) complements these results by obtaining the limiting value of such coefficients and shows that for any  $ARIMA(p, d, q)$  variable the limiting model is  $IMA(d, d)$ . In fact, this result coincides with Working's for the particular case of the random walk where an  $ARIMA(0, 1, 0)$  aggregates to an  $IMA(1, 1)$ .

<sup>2</sup>A comprehensive survey of temporal aggregation can be found in Silvestrini and Veredas (2008).

the effects of two types of temporal aggregation: time averaging and interval sampling, in terms of its implications for the properties of a random walk process. We focus on a random walk given the importance of its implications for issues such as real business cycles (Nelson and Plosser, 1982), permanent income/life cycle hypothesis (Deaton, 1987), asset prices and market efficiency hypothesis (Fama, 1970). Although the frequency of the true data generating process is unknown, we assume that it is high, and thus the natural interval is equal to or less than the actual interval of observations used for any form of analysis. We undertake our analysis within the context of asset prices that in theory should follow a martingale process and their returns a martingale difference. We also consider the implications of temporal aggregation on the variance ratio test, which has often been employed in both the macroeconomic and financial literatures. The variance ratio test, which simply requires estimation of the variance of a time series (and its  $k^{th}$  difference), has been used as an alternative to unit root tests, by testing for uncorrelated increments in areas ranging from exchange rates (Grilli and Kaminsky, 1991) to real output (Campbell and Mankiw, 1987). Yet several studies have found contradictory results using the same data that differ only by the extent of temporal aggregation, for example in testing the random walk hypothesis in the exchange rate literature (Liu and He, 1991; Yilmaz, 2003).

Thus, we pursue our investigation of the aggregated data on four fronts and can summarize our contribution as follows. First we utilize a notational framework that allows us to make a direct comparison of the properties of the aggregated data with that of the disaggregated data. The framework we use is most similar to the one used by Tiao (1972) who derives general expressions for the impact of temporal aggregation in integrated time series models. In comparison to Tiao (1972), the notation and resulting expressions which we derive are considerably simpler, although we focus specifically on the random walk case. We are able to provide exact results that may be used for inference when considering how the *degree of temporal aggregation* plays a role on the

properties of the aggregated time series. More specifically, we use our framework to derive the first and second moments of the aggregated series analytically and we are able to easily translate and compare the properties of the aggregated and disaggregated series. In doing so, we are able to determine whether a difference exists between the two, which is useful when attempting to infer something about the populations moments of the true data generating process. We focus our attention on the moments of the aggregated series rather than attempting to estimate parameters of the data generating process, which has already received prior attention in the literature.<sup>3</sup>

Second, we seek to characterize the nature of the biases that emerge from temporal aggregation. Although knowledge of the existence of the bias in the time averaging case is not something that is new (see Working, 1960, and Taylor, 2001), we make a contribution to the extant literature by demonstrating that these biases can arise as a result of the aggregation process itself, by linking the degree of aggregation directly to the magnitude of the bias.<sup>4</sup> In particular, one key insight that we provide is to demonstrate that the total bias may be decomposed into two parts:- a small sample bias, as well as a distortion that is introduced as a result of time averaging. In addition, we derive exact expressions for the autocorrelation and autocovariance terms, for any order, both for levels and general differences of the series. Consequently, we are able to infer how the degree of temporal aggregation impacts these terms explicitly.

Third, we analytically derive the implications of temporal aggregation on the Variance Ratio in the time averaging case and find that the variance of increments in the aggregated series is not linear in the sampling interval. In doing so, we are able to generalize the result found by Working (1960) in computing the variance (and covariances) of a differenced series, from first differences to

---

<sup>3</sup>See for example Zellner (1966) who suggests an iterative method for obtaining maximum likelihood estimates of parameters; Telser (1967) also presents an estimating procedure for obtaining consistent estimates through the use of the variance and autocorrelogram of the observed aggregated series. Sargan (1974) demonstrates the biases that emerge in estimates going from a continuous time process to a discrete time approximation of that process.

<sup>4</sup>The only other work of which we are aware of that resembles the approach we take, aside from Working (1960), is that of Campos, Ericsson and Hendry (1990), who utilize a similar approach to derive the implications of phase averaging data.

the  $q^{th}$  difference. Moreover, since our framework allows for an explicit formulation of the degree of temporal aggregation, our analysis provides the tools to be able to make direct comparisons of the variance ratio across different levels of temporal aggregation, as well as computing the limits of the variance ratio, both of which were not previously possible in the existing literature. We demonstrate that our results have important repercussions for variance ratio tests that have been used extensively in the asset pricing and other literatures and use a simple empirical example to illustrate the effects.

Finally, we consider the impact of temporal aggregation on Sharpe ratios. In particular, we derive an expression for the Sharpe ratio that accounts for the case where data has been time averaged, and we briefly examine the implications. Here under the maintained assumption that the true data generating process for returns is a random walk at the highest frequency, we are able to demonstrate that the value of the Sharpe ratio increases with the degree of temporal aggregation. We are also able to determine how the holding period for assets can also affect the value of the Sharpe ratio and find that it is negative. However overall, we show that the temporal aggregation effect dominates. The remainder of the paper is structured as follows. Section 2 highlights the nature of temporal aggregation and derives expressions for covariances and correlations under the two different types of aggregation that we consider. Section 3 focuses on the impact of temporal aggregation for the variance ratio. Section 4 derives results for differenced series. Section 5 examines the implications for the Sharpe ratio and the final section concludes.

## **2 Temporal Aggregation**

There are two types of temporal aggregation that we investigate within this paper. The first is that of time averaging where data is aggregated by averaging the values of a series within a non-overlapping interval. The second type of temporal aggregation that we explore within this paper is

that of interval sampling, where an aggregated dataset is created by sampling the data at a lower frequency. We begin by outlining the two approaches and the bias that they generate in what follows. The derivations of all the analytical results can be found in the technical appendix which accompanies this paper and is available on the author's websites. Consider a series  $y_t$  that follows a random walk, where increments are uncorrelated<sup>5</sup>

$$y_t = y_{t-1} + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (1)$$

where  $\varepsilon_t$  is white noise:  $E(\varepsilon_t) = 0$ ;  $Var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma_\varepsilon^2$ ;  $E(\varepsilon_t \varepsilon_s) = 0 \forall t \neq s$  and without any loss of generality we assume the initial condition to be zero,  $y_0 = 0$ . In this case, the mean of the series,  $E(y_t) = 0$ , and the  $k^{th}$  covariance and autocorrelation are simply

$$Cov(y_t, y_{t-k}) = (t-k)\sigma_\varepsilon^2; \quad Cor(y_t, y_{t-k}) \equiv \rho_k = \sqrt{\frac{t-k}{t}}. \quad (2)$$

The variance of the series is simply obtained by setting  $k = 0$  in the expression above and equals  $t\sigma_\varepsilon^2$ .<sup>6</sup>

## 2.1 Temporal Aggregation through Time Averaging

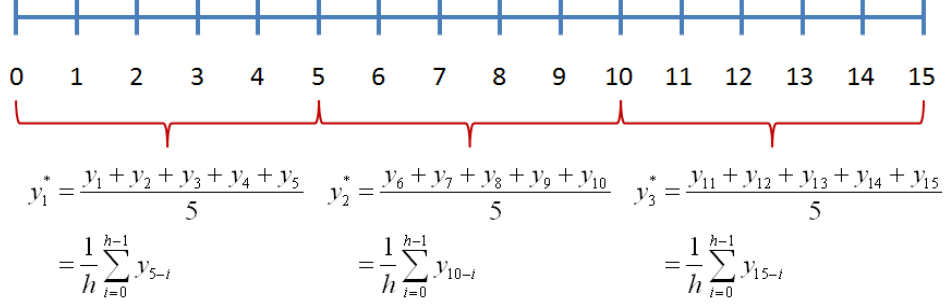
Under time averaging, a lower frequency dataset is generated by computing the average value of the data in non-overlapping intervening periods. For example, suppose that the original data generating process is at a daily frequency and that we average every five periods to obtain a weekly one, i.e.  $h = 5$  as below.

---

<sup>5</sup>The addition of a drift term is straightforward and does not alter the main results obtained in the paper.

<sup>6</sup>For ease of exposition, we adopt this convention in the remainder of the paper, where the variance can simply be computed by setting  $k = 0$  (or  $\bar{k} = 0$  in what follows in the subsequent sections) within the respective covariance term.





Let  $\tilde{t}$  represent observations from the aggregated data, where  $\tilde{t} = \frac{t}{h}$ , and  $\tilde{k}$  represent lags of the aggregated series, i.e.  $\tilde{k} = \frac{k}{h}$ . This notation allows the computation of the results to be expressed in either the ‘units of time’ of the true data generating process (henceforth DGP) and the corresponding aggregated process. An aggregated series  $y_{\tilde{t}}^*$  (in terms of the original observations) can be constructed as

$$y_{\tilde{t}}^* = \frac{1}{h} \left( y_{h(\tilde{t}-1)+1} + \dots + y_{h\tilde{t}} \right) = \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-i}.$$

The  $\tilde{k}^{th}$  lag of the aggregated series can then be represented as  $y_{\tilde{t}-\tilde{k}}^* = \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-h\tilde{k}-i} = \frac{1}{h} \sum_{i=0}^{h-1} y_{h(\tilde{t}-\tilde{k})-i}$ .

When examining the properties of the aggregated series  $y_{\tilde{t}}^*$ , where we set  $\tilde{k} = 0$ , we can show that the mean,  $E(y_{\tilde{t}}^*) = E[y_{h\tilde{t}}] = 0$ , and the  $\tilde{k}^{th}$  covariance equals

$$Cov(y_{\tilde{t}}^*, y_{\tilde{t}-\tilde{k}}^*) = \sigma_{\varepsilon}^2 \left[ h(\tilde{t} - \tilde{k}) - (h-1) + \frac{(h-1)(2h-1)}{6h} \right].$$

When  $h = 1$ , then the expressions for the variance and covariance above reduces to  $Var(y_{\tilde{t}}^*) = t\sigma_{\varepsilon}^2$  and  $(t-k)\sigma_{\varepsilon}^2$  - what we obtained before with no temporal aggregation. Similarly, the correlation between  $y_{\tilde{t}}^*$  and  $y_{\tilde{t}-\tilde{k}}^*$  can thus be expressed as

$$Cor(y_{\tilde{t}}^*, y_{\tilde{t}-\tilde{k}}^*) \equiv \rho_{\tilde{k}}^* = \frac{Cov(y_{\tilde{t}}^*, y_{\tilde{t}-\tilde{k}}^*)}{(Var(y_{\tilde{t}}^*))^{0.5} (Var(y_{\tilde{t}-\tilde{k}}^*))^{0.5}} = \sqrt{\frac{h(\tilde{t} - \tilde{k}) + D}{h\tilde{t} + D}},$$

where  $D(h) \equiv \frac{(h-1)(2h-1)}{6h} - (h-1) = -\frac{(h-1)}{6h} [1 + 4h] \leq 0$  since  $h \geq 1$ . Since  $t = h\tilde{t}$ , note that  $\lim_{t \rightarrow \infty} \rho_{\tilde{k}}^* = 1$  and when  $h = 1$ ,  $\rho_{\tilde{k}}^* = \rho_k$ . Similarly, expressions for the variance and covariance

above can be expressed in terms of the distortionary term,  $D(h)$ , as  $Var(y_t^*) = [h\tilde{t} + D] \sigma_\varepsilon^2$  and  $Cov(y_t^*, y_{t-\tilde{k}}^*) = [h(\tilde{t} - \tilde{k}) + D] \sigma_\varepsilon^2$  respectively. When examining the total distortion in the time averaging case, we note that  $\frac{\partial D}{\partial h} = -\left(\frac{4h^2+1}{6h^2}\right) < 0$  which would indicate that as the distance between observations,  $h$ , increases, the magnitude of the distortion is increasing in absolute terms. A final noteworthy point is that the presence of the distortion term above introduces a distinct difference between an aggregated series and a disaggregated series. Given our framework, we can examine the difference between the two, conditioning either on a specific time period, or on a specific order of autocorrelation. For example, conditioning on a specific period of time, we are able to compute the difference between the value for the correlation coefficient for the temporally aggregated series less the corresponding (true) value for the disaggregated series as

$$b_A^\rho = \rho_{\tilde{k}}^* - \rho_k = \sqrt{\frac{t-k+D}{t+D}} - \sqrt{\frac{t-k}{t}}, \quad (3)$$

which is asymptotically zero as  $t \rightarrow \infty$ . However, we could instead condition on the specific order of autocorrelation, and the corresponding difference for the correlation coefficient between the aggregated series and its disaggregated counterpart would then equal

$$b_A^\rho = \rho_{\tilde{k}}^* - \rho_{\tilde{k}} = \sqrt{\frac{t-k+D}{t+D}} - \sqrt{\frac{t-k/h}{t}}. \quad (4)$$

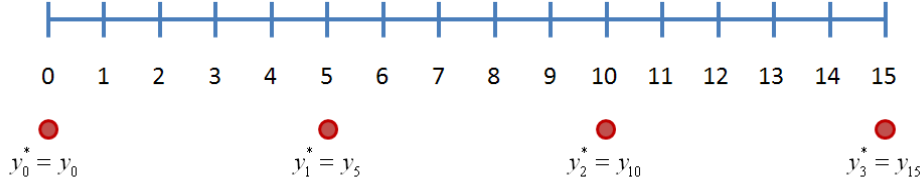
For the covariance terms, a similar difference between an aggregated series and its disaggregated counterpart, may be written as

$$b_A^C = Cov(y_t^*, y_{t-\tilde{k}}^*) - Cov(y_t, y_{t-k}) = [h(\tilde{t} - \tilde{k}) + D] \sigma_\varepsilon^2 - (t-k) \sigma_\varepsilon^2 = D\sigma_\varepsilon^2 \quad (5)$$

when one conditions on a specific period of time. In the case where we condition on a specific value of  $\tilde{k}$ , we would instead obtain  $b_A^C = Cov(y_t^*, y_{t-\tilde{k}}^*) - Cov(y_t, y_{t-\tilde{k}}) = \sigma_\varepsilon^2 [D - (\frac{h-1}{h})k]$ . In examining the expressions above, we note that these differences equal to zero when  $h = 1$ , but will

differ from zero in general when  $h > 1$ .

## 2.2 Temporal Aggregation through Interval Sampling



Under the alternative form of temporal aggregation, prices are obtained by taking every  $h$ -th observation of the series. Suppose that a dataset is generated by sampling the process given in equation (1) every  $h$  observations. This could occur when a lower frequency dataset is created from a higher frequency one by using end of period observations or beginning of period observations. For example, suppose that the original data generating process is at a daily frequency and that we use beginning of period values every five periods to obtain a weekly series, i.e.  $h = 5$  as above. The  $\tilde{k}^{th}$  lag of the aggregated series can be represented in terms of the original observations as  $y_{\tilde{t}-\tilde{k}}^* = y_{h\tilde{t}-h\tilde{k}} = y_{h(\tilde{t}-\tilde{k})}$ . As earlier under the case of time averaging, the mean of the aggregated series  $E(y_{\tilde{t}}^*) = E(y_{h\tilde{t}}) = 0$ , whilst the  $\tilde{k}^{th}$  covariance

$$Cov(y_{\tilde{t}}^*, y_{\tilde{t}-\tilde{k}}^*) = (\tilde{t} - \tilde{k}) h \sigma_{\varepsilon}^2, \quad (6)$$

and whose  $\tilde{k}^{th}$  autocorrelation is  $Cor(y_{\tilde{t}}^*, y_{\tilde{t}-\tilde{k}}^*) \equiv \rho_{\tilde{k}}^* = \sqrt{\frac{(\tilde{t}-\tilde{k})}{\tilde{t}}}$ . Based on the true DGP, the true variance, covariance and correlation are given in equation (2). When examining the difference between the properties of an aggregated series to that of a disaggregated series, like that performed earlier, we find that there is no difference for the correlation coefficient in the interval sampling case, since  $b_I^{\rho} = \rho_{\tilde{k}}^* - \rho_k = \sqrt{\frac{(t-k)/h}{t/h}} - \sqrt{\frac{t-k}{t}} = 0$ . Similarly, it is possible to see from comparing equation (6) above to (2) that there will be no distortion in the variance or covariance terms and these terms can be consistently estimated.

However, when we once again condition on a specific value of  $\tilde{k}$ , we find that the aggregation process yields a smaller sample than what was available for the higher frequency process, and this then leads to a difference between the aggregated and disaggregated series. For the differences in the correlation coefficient, this equals

$$b_I^{\rho} (h, \tilde{k}) = \rho_{\tilde{k}}^* - \rho_{\tilde{k}} = \sqrt{\frac{(t-k)}{t}} - \sqrt{\frac{t-k/h}{t}} \neq 0. \quad (7)$$

For the variance and covariance terms it equals

$$b_I^C = Cov(y_t^*, y_{t-\tilde{k}}^*) - Cov(y_t, y_{t-k}) = (\tilde{t} - \tilde{k}) \sigma_{\varepsilon}^2 - (t-k) \sigma_{\varepsilon}^2 = -\left(\frac{h-1}{h}\right) (t-k) \sigma_{\varepsilon}^2. \quad (8)$$

Two final points merit a mention. The first pertains to the issue of small sample biases that may arise from the temporal aggregation process. The small sample bias is reflected in the term  $-\left(\frac{h-1}{h}\right)$  above and in the corresponding expression in the time averaging case. The emergence of this bias is due to the loss of information arising from the aggregation process. The use of the term ‘small sample bias’ here differs to its customary usage, since it is related to  $h$ , the degree of aggregation, rather than the conventional small sample bias that arises from a smaller value of  $t$ .

The second observation we make is that as the distance between observations,  $h$ , increases, the magnitude of the difference between the aggregated and disaggregated series increases in absolute terms, making the underlying process appear to be more stationary than it truly is. It is also interesting to note that these small sample biases disappear asymptotically, or if  $h = 1$ . Hence, in the case of interval sampling, any bias that we might observe would arise due to a small sample problem, since when we condition on  $\tilde{k}$ , any difference that remain relates to the the loss of information resulting from temporal aggregation. However, as we find when we condition on a specific period of time, parameter estimates will still be consistent when estimating equation (1).<sup>7</sup>

---

<sup>7</sup>The intuition behind why parameter estimates will be consistent is as follows. The only difference between the original data generated from equation (1) and the data obtained as a result of temporal aggregation, when the form

### 2.3 Decomposition of the Distortion

In examining the two types of temporal aggregation, we can see by comparing equations (8) and (5), that for  $k > 0$  and  $h \geq 1$ , that the absolute value of the differences between the aggregated series and the disaggregated one should be larger under time averaging relative to the interval sampling. It is also possible to show this by examining the difference in the correlation coefficient. Doing this comparison allows us to think of the difference in these properties in terms of a pure “small sample” bias arising from the loss of observations, as in the interval sampling case, as well as a distortionary component that is introduced in the time averaging case. Notice that equation (3) - (5) contain the term  $D$ , which can be decomposed into a “small sample” bias as well as a distortionary component. If we define the loss of information arising from the aggregation process as  $B^s \equiv -\frac{(h-1)}{h}$ , which was observed in the interval sampling case for the variance and co-variance, then we can express  $D(h)$  in terms of a small sample component and a distortionary component as follows

$$D = B^s \frac{(1 + 4h)}{6} = B^s B^d. \quad (9)$$

Thus  $B^d \equiv \frac{1}{6}(1 + 4h)$  represents the distortionary part arising as a result of time averaging. Thus for  $h > 1$ , the total distortion in the time averaging case will be larger in absolute value than what appears in the interval sampling case since  $B^d > 1$ .

### 3 Variance Ratio Test

In this section we examine the effect that time averaging has on both the variance of increments of the aggregated process and the calculation of the Variance Ratio. The true DGP we are considering exhibits uncorrelated increments  $\Delta y_t = \varepsilon_t$ , where  $\Delta$  is the first difference operator, and  $Var(\Delta y_t) =$  

---

 of aggregation is interval sampling, is merely the number of data points available to estimate the coefficients. As such, the smaller (aggregated) data will still yield consistent estimates, although it may have larger standard errors.

$\sigma_\varepsilon^2$ .<sup>8</sup> The  $k$ -difference of  $y_t$ ,  $\Delta_k y_t = \sum_{i=0}^{k-1} \varepsilon_{t-i}$ , and therefore  $Var(\Delta_k y_t) = k\sigma_\varepsilon^2$ . The ratio of the

variance of a  $k$ -period interval of a random walk to  $k$  times the variance of a one-period interval,

the Variance Ratio,  $VR(k) = \frac{Var(\Delta_k y_t)}{kVar(\Delta y_t)}$ , should be one, or statistically indistinguishable from one.<sup>9</sup>

This property has been used to derive single-tests of unity for the Variance Ratio for individual values of  $k$ , consistent under both homoskedasticity and heteroskedasticity (see Lo and MacKinlay, 1988) as well as multiple Variance-Ratio tests (see Chow and Denning, 1993).

Monte Carlo experiments show that these test statistics are more powerful than alternative unit root tests such as DF, PP or Box-Pierce when testing for random walks (see Lo and MacKinlay, 1989; Liu and He, 1991; and Chow and Denning, 1993). As such, variance ratios have been applied to test the null hypothesis of a random walk for returns on equity, exchange rates and interest rates in developed as well as emerging markets (see e.g. Liu and He, 1991; Fong, Koh, and Ouliaris, 1997; Yilmaz, 2003; Belaire-Franch and Opong, 2005).<sup>10</sup> Many of these studies suggest a rejection of the martingale property of financial returns, and sometimes indicate predictability with respect to their own past. However, studies using similar data at different frequencies find contradictory results (which we highlight a little later).

To illustrate the effect of temporal aggregation, we compute the variance of the  $\tilde{k}$ -difference

$$Var(\Delta_{\tilde{k}} y_t^*) = \left( h\tilde{k} + 2D + (h - 1) \right) \sigma_\varepsilon^2 \quad (10)$$

Note that for  $\tilde{k} = 1, h = 1$  we obtain the result for the original series,  $Var(\Delta y_t^*) = Var(\Delta y_t) = \sigma_\varepsilon^2$ .

Our result allows for an easy computation of the variance of the  $k$ -difference of the series and can nest the result in Working (1960) for the first difference of the series, or returns of the aggregated

---

<sup>8</sup>One can think of  $y_t$  as log prices ( $\log P_t$ ) and therefore  $\Delta y_t$  as the return series.

<sup>9</sup>The Variance Ratio will still approach one asymptotically even in the case of dependent but uncorrelated increments. This holds because the variance of the sum of uncorrelated increments equals the sum of the variances.

<sup>10</sup>Variance ratios have also been used to measure the persistence of real output (see e.g. Campbell and Mankiw, 1987; Cochrane, 1988; Cecchetti and Lam, 1994).

series. In particular, for  $\tilde{k} = 1$ ,  $Var(\Delta y_t^*) = \left(1 + \frac{(2h-1)(h-1)}{3h}\right) \sigma_\varepsilon^2 = \frac{2h^2+1}{3h} \sigma_\varepsilon^2$ . We also generalize our result by obtaining a general expression of the Variance Ratio for the aggregated series for any value of  $\tilde{k}$  and  $h$ :

$$VR(\tilde{k}) = \frac{Var(\Delta_{\tilde{k}} y_t^*)}{\tilde{k} Var(\Delta y_t^*)} = \frac{h\tilde{k} + 2D + (h-1)}{\tilde{k}(2D + (2h-1))}. \quad (11)$$

The Variance Ratio is therefore non-constant, since it is a function of  $h$  and  $\tilde{k}$ . By examining equation (11) above, we can determine how the Variance Ratio changes as either  $\tilde{k}$  changes or as  $h$ , the degree of temporal aggregation changes. For example, for a given value of  $\tilde{k}$ , we can compute the asymptotic behavior with respect to  $h$ . This yields  $\lim_{h \rightarrow \infty} VR(\tilde{k}) = \frac{3}{2} \left[1 - \frac{1}{3\tilde{k}}\right]$ . From this expression, we see a random walk process is bounded by 1, and it has an asymptote at 1.5 as the level of aggregation  $h$  and  $\tilde{k}$  grow without bound (see Figure 1).<sup>11</sup> A similar pattern for the impulse response function (*irf*) emerges given its close relationship with the variance ratio. Campbell and Mankiw (1989) derive a lower bound on the value of *irf* or the infinite sum of the moving average coefficients as  $irf = \sqrt{\frac{v}{var(\varepsilon_t)/var(\Delta y_t)}} = \sqrt{v}$  (for a random walk process), where  $v = \lim_{\tilde{k} \rightarrow \infty} VR(\tilde{k}) = \frac{h}{\frac{1}{3h}(2h-1)(h-1)+1}$ . The value of *irf* has a lower bound at 1.225 as  $h$  grows indefinitely.<sup>12</sup>

The results above imply that the Variance Ratio test typically used, would be invalid if the frequency of the data does not match the one of the true DGP. For example, assuming that the ‘true’ higher frequency process follows a random walk, the Variance Ratio in (11) applied to a lower frequency dataset would be biased upwards, which would increase the type I error of the test. This can easily be seen from the following example. Suppose that we were to assume that log prices,  $y_t$ , followed

<sup>11</sup>The asymptotic value of 1.5 for the Variance Ratio and of 1.26 for the impulse response function was previously mentioned in Rossana and Seater (1995). However, they did not link the degree of aggregation,  $h$ , to the bounds imposed on the Variance Ratio. Our expression allows us to show how the Variance Ratio will behave for any  $k^{th}$  difference as well as when the degree of aggregation,  $h$ , varies.

<sup>12</sup>An alternative but equivalent way of computing the *irf* is one plus the sum of the autocorrelation coefficients, which for the case of  $\tilde{k} = 1$  is equal to  $1 + \frac{h^2-1}{4h^2+2}$ , and for  $h \rightarrow \infty$  is equal to 1.25.

a random walk at the daily frequency and therefore returns ( $\Delta y_t$ ) were white noise. The use of weekly data ( $h = 5$ ) would generate the following values for the variance ratio at lags  $\tilde{k} = 2, 4, 8,$  and  $16$  :  $VR(2) = 1.23, VR(4) = 1.35, VR(8) = 1.41, VR(16) = 1.44$ . Liu and He (1991, table II, pp. 777) find the following average values when computing variance ratio tests for the dollar exchange rates vis a vis the Deutschmark, the Yen and the Pound Sterling, where the null of a random walk was rejected using weekly data:  $VR(2) = 1.04, VR(4) = 1.13, VR(8) = 1.23, VR(16) = 1.34$ . Interestingly enough, Yilmaz (2003, table 4) using daily frequency data for the same set of currencies found variance ratio tests that were much lower and closer to one, though all greater than one, than the ones in Liu and He (1991).

We complement those results in the literature for the dollar/sterling exchange rate by, first, updating the data, and secondly, by using both time average and sampling aggregated data. The beginning of the sample is the same as in Liu and He, that is, August 7, 1974, and extending it until May 13, 2011. We use daily data available from the FRED database as the highest frequency, as well as their time averaged and interval sampled data, available at the weekly and monthly frequency. The results in Table 1 Panel A show that variance ratios for the time average aggregated data are always significant, larger than one and larger than the  $VR$  of the daily series. For the interval sampling case this result does not hold.<sup>13</sup>

Our result has consequences, not only for the rejection or not of the random walk hypothesis, but also for the interpretation of the Variance Ratio. We can note from equation (11) that  $Var(\Delta_{\tilde{k}} y_t^*)$  does not grow linearly with  $\tilde{k}h$ . Values of the Variance Ratio above one suggest positive serial correlation given that the  $VR(\tilde{k})$  can be interpreted as the weighted sum of the first  $\tilde{k} - 1$  autocorrelation coefficients. For example in the case of exchange rate returns, the bias induced by temporal aggregation will favour the undershooting phenomenon (positive autocorrelation) relative

---

<sup>13</sup>We have run Monte Carlo experiments where these results are ratified using random walks as DGP and different levels and methods of aggregation. Results are not reported for space consideration but available upon request.



to the overshooting theory (negative autocorrelation) (see Liu and He, 1991; and Huizinga, 1987).

## 4 Autocorrelations of Differenced Series

We now focus our attention on the covariance and autocorrelations of a  $\tilde{k}$ -differenced sequence. In order to do so, we need to introduce a new parameter  $\tilde{q}$ , which represents the autocovariance order for the  $\tilde{k}$ -differenced aggregated series. In comparing the results that follow here to those of Working (1960), we note that the results derived in Working (1960) are for the case of the first-order autocovariance ( $\tilde{q} = 1$ ) for the first-differenced ( $\tilde{k} = 1$ ) aggregated series. In terms of our notation, this would amount to setting the values  $\tilde{k} = 1, \tilde{q} = 1$  within the autocovariance term  $Cov(\Delta_{\tilde{k}}y_t^*, \Delta_{\tilde{k}}y_{t-\tilde{q}}^*)$  i.e.  $Cov(\Delta y_t^*, \Delta y_{t-1}^*)$ . Given that  $\Delta_{\tilde{k}}y_t^*, \Delta_{\tilde{k}}y_{t-\tilde{q}}^*$  are only sums of white noise processes, the expression for the covariance is derived from the product of  $\Delta_{\tilde{k}}y_t^*$  and  $\Delta_{\tilde{k}}y_{t-\tilde{q}}^*$ , which will have values different from zero as long as the following condition is met:  $h(1 + \tilde{k} - \tilde{q}) \geq 2$ . This condition holds when  $\tilde{k} \geq \tilde{q}$  given that  $h > 1$ . The covariance has two expressions depending on the values of  $\tilde{k}$  and  $\tilde{q}$ :

$$Cov(\Delta_{\tilde{k}}y_t^*, \Delta_{\tilde{k}}y_{t-\tilde{q}}^*) = \begin{cases} \frac{1}{h^2} \left[ \sum_{i=1}^{h-1} i(h-i)\sigma_\varepsilon^2 \right] = \frac{(h^2-1)}{6h}\sigma_\varepsilon^2 & \text{if } \tilde{q} = \tilde{k} \\ h(\tilde{k} - \tilde{q})\sigma_\varepsilon^2 & \text{if } \tilde{k} > \tilde{q} \end{cases}. \quad (12)$$

When comparing the expression above to those that Working (1960) derived, we note that our expression is more general and nests the one obtained by Working who examined the case where  $\tilde{q} = \tilde{k} = 1$ . In his particular case, the autocovariance terms are zero,  $\forall \tilde{q} \geq 2$ . However, for the case that  $\tilde{k} > \tilde{q}$ , no distortion exists in the expression for the covariance.<sup>14</sup>

However, the correlation coefficient will include a distortion that arises from the variance term.

For the case of  $\tilde{k} = 1$  (and therefore  $\tilde{q} = 1$ , which is the specific case that Working examined), we

---

<sup>14</sup>This is because in the true DGP, the  $q^{th}$  order autocovariance term  $Cov(\Delta_k y_t, \Delta_k y_{t-q}) = (k - q)\sigma_\varepsilon^2$ , where we use the notational transformation to aggregated data  $(k - q) = h(\tilde{k} - \tilde{q})$ .

derive the following expression for autocorrelation:

$$Cor(\Delta y_t^*, \Delta y_{t-1}^*) = \frac{Cov(\Delta y_t^*, \Delta y_{t-1}^*)}{\left(Var(\Delta y_t^*)\right)^{0.5} \left(Var(\Delta y_{t-1}^*)\right)^{0.5}} = \frac{\frac{h^2-1}{6h}}{\frac{2h^2+1}{3h}} = \frac{h^2-1}{4h^2+2}. \quad (13)$$

When  $\tilde{k} > 1$  we once again have to consider two scenarios. In the first scenario, we consider the case where  $\tilde{k} = \tilde{q}$ , and we obtain:

$$\begin{aligned} Cor(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*) &= \frac{Cov(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*)}{\left(Var(\Delta_{\tilde{k}} y_t^*)\right)^{0.5} \left(Var(\Delta_{\tilde{k}} y_{t-\tilde{q}}^*)\right)^{0.5}} \\ &= \frac{\frac{h^2-1}{6h}}{h(\tilde{k}-1) + 1 + \frac{(2h-1)(h-1)}{3h}} = \frac{\frac{h^2-1}{6h}}{h\tilde{k} + 2D + (h-1)} \text{ if } \tilde{q} = \tilde{k}. \end{aligned} \quad (14)$$

Given that  $\tilde{q} = \tilde{k}$ , and considering any values of  $\{\tilde{k}, \tilde{q}\}$  for a given amount of temporal aggregation  $h$ , the distortion in the autocorrelation coefficient differs from Working's due to the distortion in the variance as can be seen by the terms in the denominator. Although this term is positive, it is worth noting that the distortion  $D$  is purely a function of the degree of temporal aggregation,  $h$ , i.e.  $D(h)$ . Consequently, we can see that the effect of the distortion term in the autocorrelation coefficient gets smaller as we increase  $\tilde{k}$ , and in the limit as  $\tilde{k} \rightarrow \infty$ , the term  $h\tilde{k}$  dominates the denominator (see Figure 2).

In the second scenario, we examine the case where  $\tilde{k} > \tilde{q}$ , and obtain the following for the autocorrelation expression:

$$Cor(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*) = \frac{h(\tilde{k}-\tilde{q})}{h\tilde{k} - (h-1) + \frac{(2h-1)(h-1)}{3h}} \text{ if } \tilde{k} > \tilde{q}. \quad (15)$$

The distortion in this correlation coefficient once again arises due to the distortion in the variance of the aggregated series given that  $Cor(\Delta_k y_t, \Delta_k y_{t-q}) = \frac{k-q}{k} = \frac{h(\tilde{k}-\tilde{q})}{h\tilde{k}}$  in the true DGP. For both equations (14) and (15), it is easy to note that the distortion arises from the second and third

terms in the denominator  $-(h-1) + \frac{(2h-1)(h-1)}{3h} = 2D + (h-1)$ . This term is negative although smaller than  $h\tilde{k}$  in absolute value, making the correlation appear larger than it should be. This implies that, as the level of temporal aggregation ( $h$ ) increases, the distortion in the autocorrelation coefficient increases, which is consistent with the previous results.

Panel B in Table 1 displays the autocorrelation coefficient for the currency return series ( $Cor(\Delta y_t^*, \Delta y_{t-\tilde{q}}^*)$ ) using both types of aggregation. It is worth noting that the autocorrelation increases with the level of aggregation for the time averaging case but not for the interval sampling one. These results are reflected in the autocorrelation values of order  $q$  for the  $k$ -differenced exchange rate series reported in Table 2. The values of the daily exchange rate very closely resemble those of a theoretical random walk. In addition, it is interesting to note that the autocorrelations for the interval sampling case do not vary with the level of aggregation. For the time averaging case, the values of the autocorrelation coefficient are higher than the theoretical values and this appears to be consistent with our results above.

### **Filtering the Effects of Temporal Aggregation**

The analytical expressions reported in (12)-(15) and the results reported in Table 2 reveal the two different effects of time averaging. First, the autocorrelation coefficients induced by the MA terms created by overlapping ‘return’ series are higher than those predicted theoretically. Second, temporal aggregation induces additional moving average terms and this results in higher order autocorrelation. For instance, in Table 2, in the case of ‘weekly’ aggregation ( $h = 5$ ), the first and second order autocorrelation coefficients of two period ‘returns’ ( $k = 2$ ,  $q = 1$  and  $q = 2$ ) are 0.591 and 0.109 (0.595, 0.095 if we use expressions (14) and (15)), respectively, which differ from the theoretical values of a random walk of 0.5 and 0, respectively.

In principle, if the researcher knew the true frequency of the DGP, the results obtained in the previous section could be used to filter the aggregated series and retrieve the properties of the

original series. In the example above, the second-differenced series has in theory an MA(1) process with first order autocorrelation of 0.5, which implies an MA term with coefficient equal to 1, i.e.  $\Delta_2 y_t = (1 + L)\varepsilon_t$ . However, the aggregated series turns out to be an MA(2) with coefficients  $\theta_1$  and  $\theta_2$ ,  $\Delta_2 \tilde{y}_t^* = (1 + \theta_1 L + \theta_2 L^2)\varepsilon_t$ . Given that we know the value of the two autocorrelation coefficients based on equations (14) - (15) are (0.595, 0.095), we could use the expression of the autocorrelation function of an  $MA(q)$  process<sup>15</sup> to obtain the values of  $\theta_1$  and  $\theta_2$ . These values would then be used to apply the filter (a Taylor expansion of)  $\frac{1+L}{1+\theta_1 L+\theta_2 L^2}$  to the series  $\Delta_2 \tilde{y}_t^*$ . In general, this procedure applies generally for any value of  $h, q$ , and  $k$ , and allows a researcher who has some prior about the natural frequency of the DGP to infer something about its true properties.

## 5 Sharpe Ratio

Sharpe ratios play an important role in finance in assessing whether investments are safer in the long run compared to the short run (see Siegel, 1988, Lettau and Ludvigson, 2010). If the standard deviation of returns grows more quickly than its mean, the Sharpe ratio grows slower than the square root of the horizon. From this, it may appear reasonable to conclude that stocks are safer in the short run than they are in the long run, given the mean-variance tradeoff. However, we illustrate below how the temporal aggregation results obtained above may affect the analysis and interpretation of the Sharpe ratio. Moreover, we use this example to illustrate how aggregation has an effect on both the ‘absolute’ value of variances within a given frequency and the ‘relative’ value across frequencies.

We begin by maintaining the assumption that, at the highest frequency, the DGP is a random walk with *iid* errors but now we assume that the process also include a drift term,  $\mu$  :

---

<sup>15</sup>The ACF of an MA( $q$ ) process can be written as  $\rho(s) = \frac{\sum_{j=0}^{q-s} \theta_j \theta_{j+s}}{\sum_{j=0}^q \theta_j^2}$  with  $\theta_0 = 1$ .

$$\Delta y_t = \mu + \varepsilon_t, \quad t = 1, 2, \dots, T \quad (16)$$

where  $E(\Delta y_t) = \mu$ ,  $Var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma_\varepsilon^2$ . In order to compare the Sharpe ratio at different horizons,  $S_k$ , we need to utilize expressions for the mean and variance of a  $k$ -differenced series, i.e.  $E(\Delta_k y_t) = k\mu$ , and  $Var(\Delta_k y_t) = k\sigma_\varepsilon^2$ . For illustrative purposes let us assume that  $y_t$  follows process (16) at the ‘daily’ frequency and supposed that we wish to compare an investment at horizons 1, 3, 5 and 10 years ( $k = 240, 720, 1200, 2400$ ). In this case, as pointed out by Lo (2002), the Sharpe ratio divided by the square root of the horizon is constant for all  $k$  :

$$\zeta \equiv \frac{S_k}{\sqrt{k}} = \frac{k\mu}{\sqrt{k\sigma_\varepsilon^2}\sqrt{k}} = \frac{\mu}{\sigma_\varepsilon} \quad (17)$$

From this, one may conclude that if  $y_t$  is a random walk with drift, then investments at different horizons are ‘equally safe’. However, this conclusion changes when we use aggregated data, and two artifacts of aggregation are worth mentioning. First, the numerator above in (17) grows linearly with  $k = \tilde{k}h$ . For example, if the level of aggregation  $h$  is annual ( $h = 240$ ), then the expected value of  $\Delta y_t^*$  at the  $\tilde{k}^{th}$  horizon ahead (e.g.  $\tilde{k} = 1, 3, 5$ ) will be  $E(\Delta_{\tilde{k}} y_t^*) = \tilde{k}h\mu = k\mu$ , which coincides with numerator in (17). On the other hand, when looking at the denominator of the Sharpe ratio, the variance of the differenced series  $Var(\Delta_{\tilde{k}} y_t^*)$ , is given by the expression in (10), and this expression does not grow linearly with  $\tilde{k}h$ . Hence the Sharpe ratio for the aggregated series is not constant, and instead may be written as:

$$\begin{aligned} \zeta^* &\equiv \frac{S_{\tilde{k}h}^*}{\sqrt{\tilde{k}h}} = \frac{\tilde{k}h\mu}{\sqrt{Var(\Delta_{\tilde{k}} y_t^*)}\sqrt{\tilde{k}h}} = \frac{\mu}{\sigma_\varepsilon} \frac{\tilde{k}h}{\sqrt{\tilde{k}h}\sqrt{h(\tilde{k}-1) + 1 + \frac{(2h-1)(h-1)}{3h}}} \\ &= \frac{\mu}{\sigma_\varepsilon} \left( \frac{\tilde{k}h}{\sqrt{\tilde{k}h}\sqrt{h\tilde{k} - h + 1 + \frac{(2h-1)(h-1)}{3h}}} \right) \end{aligned} \quad (18)$$

Given the expression above, we note three things.

**Proposition 1** *Under the maintained assumption that at the highest frequency, the data generating process is random walk with iid errors, and  $\forall h, k \geq 1$ , we propose that:*

- (a) *For a given level of aggregation,  $h$ ,  $\frac{d\zeta^*}{dk} < 0$*
- (b) *For a given level of  $\tilde{k}$ ,  $\frac{d\zeta^*}{dh} > 0$*
- (c) *When considering joint increases in  $h$  and  $\tilde{k}$ ,  $\frac{d\zeta^*}{d(h\tilde{k})} > 0$ .*

First, consider the implications of proposition 1(a). It would indicate that for a given level of aggregation,  $h$ , and in the absence of appropriate corrections as suggested by Lo (2002), the Sharpe ratio divided by the square root of the horizon,  $\zeta^*$ , decreases with  $\tilde{k}$ . This would have the effect of making short-run investments appear to be ‘safer’ than long run investments and this feature has already been noted empirically by some (see for example Cochrane, 2001, p.412).<sup>16</sup>

Second, with regards to proposition 1(b). Given the earlier results from the variance and covariance expressions, we know that the modified Sharpe ratio,  $\zeta^*(h)$  is an increasing function of  $h$ . Thus conditioning on  $\tilde{k}$ , we should see the modified Sharpe ratio increase as we aggregate data for a given holding period,  $\tilde{k}$ . This can be seen in the second term in expression (18), which is an increasing function of  $h$ , and hence larger than 1 for  $h > 1$ . This means that expression (18) is larger than (17) making the investment ‘more profitable’ per unit of risk when using temporally aggregated data as compared to observations from the true frequency of the DGP.

Finally, proposition 1(c) considers the case when  $h$  and  $\tilde{k}$  jointly increase. In looking at these comparisons, it is possible to show that  $\frac{d\zeta^*}{d(h\tilde{k})} > 0$ . Hence aggregated data, even at comparable holding period horizons, should yield a higher value for  $\zeta^*$ .

In order to verify and validate these results, we compute the modified Sharpe ratio for different levels of temporal aggregation as a simple empirical illustration. We use the price at close of the S&P 500 to construct daily net returns between 1970 - 2013. We then compute  $\zeta^*$  at 1, 3, 5, and 10 year horizons for daily, monthly, quarterly and annually time averaged data. The results are reported in figure 3 and table 3.

---

<sup>16</sup>Note that  $\frac{h\tilde{k}}{\sqrt{h\tilde{k}}\sqrt{h\tilde{k}-h+1+\frac{(2h-1)(h-1)}{3h}}} > 1$  because  $h\tilde{k} - h + 1 + \frac{(2h-1)(h-1)}{3h} < h\tilde{k}$ .

Figure 3 depicts the results for daily data. As can be seen in the figure 3, when we increase the holding period horizon  $\tilde{k}$ , the value of the modified Sharpe ratio on the whole declines. We can also see the decline in  $\zeta^*$  for  $h > 1$ , i.e. the monthly, quarterly and annual data in panel B of table 3. For example when using the annual time averaged data, the value of the modified Sharpe ratio declines from 0.355 down to 0.278, and this is consistent with proposition 1(a). Thus for a given frequency of the data, one might reasonably conclude that short term holding periods were ‘safer’ than long term holding periods.

When we instead condition on a particular holding period horizon, we find that the value of the modified Sharpe ratio increases with the degree of temporal aggregation. For example at a holding period horizon of 1 year, the value increases from 0.025 to 0.355, and this is consistent with proposition 1(b). However, one should note that in going across the rows of panel B, we are really comparing different levels of aggregations at comparable holding period horizons, i.e.  $h\tilde{k} = \tau$ , where  $\tau$  is a constant (for example one year). This will be the case whenever we examine the modified Sharpe ratio for  $\tau$  holding periods for daily versus monthly or quarterly aggregated data. Here, for the same comparable holding period, e.g. one year ( $\tau = 1$ ), when we use daily data ( $h = 1$ ) and evaluate a one year holding period, then  $\tilde{k} = 240$ . When we look at the equivalent number for the monthly data ( $h = 20$ ), there  $\tilde{k} = 12$ . So the numbers in the rows of panel B really depict the situation where  $\partial h > 0$  and  $\partial k < 0$ . However, given the effect above for proposition 1(a), since the numbers are increasing across the rows, the only way that may happen is if  $\frac{d\zeta^*}{dh} > 0$ .

Finally, when we may examine how the modified Sharpe ratio changes based on the joint increase in  $h$  and  $k$  by (loosely) looking at the main diagonal in the results in panel B. There we see that  $\zeta^*$  increases from 0.025 ( $h = 1$ , 1 year holding period) to 0.278 ( $h = 240$ , 10 year holding period), which verifies proposition 1(c). In addition, the result here would seem to indicate that in terms of overall effects, the temporal aggregation effect dominates the holding period effect, i.e.  $\frac{d\zeta^*}{dh} > \frac{d\zeta^*}{dk}$ .

## 6 Conclusions

This paper has sought to understand the implications of temporally aggregating data when the underlying data generating process contains a random walk. As such, the findings in this paper are relevant for a large number of literatures where some fundamental process driving a series may contain a random walk. We have examined the implications here within the context of the literature on exchange rates and asset prices. We do this under two scenarios, one where the data is time averaged, and the other where the data is interval sampled.

Our key findings are as follows. We are able to characterize the difference between the aggregated and disaggregated series in terms of the autocorrelation and autocovariance functions analytically under both time averaging and interval sampling. As such, we can attribute the difference that arises between the two series to two components: a distortion that arises out of time averaging the data, as well as a small sample bias, resulting from the loss of information due to temporal aggregation, and this occurs under both time averaging and interval sampling. We derive a general analytical expression for the variance and covariance, *of any order*, for both the level and differences of an aggregated series, and this allows us to generalize the findings of Working (1960). The results here could, in principle, be used to filter the aggregated series and retrieve the original properties of the underlying series if we knew the true frequency of the data generating process.

In addition, we obtain a generalized expression for the Variance Ratio at any lag value, and note that the Variance Ratio is a function of the degree of aggregation as well as the autocorrelation lag, and that it differs from one. In doing so, we are able to link the degree of aggregation directly to value of the Variance Ratio, and in doing so, bound the set of values that the Variance Ratio can take for a process that follows a random walk.

Finally, we also examine the implications of temporal aggregation for the Sharpe ratio. We demonstrate that the use of temporally aggregated data, that has been time averaged, will increase the



value of the Sharpe ratio relative to the use of Sharpe ratios computed using higher frequency data, given that the higher frequency process follows a random walk. This would suggest that returns have a higher value for the Sharpe ratio when using annual data, as compared to the same number computed using higher frequency data. We also derive implications for how the holding period may impact the Sharpe ratio within this framework, and find that it tends to decrease the Sharpe ratio. Overall, the results we present here are relevant for anyone working with time series data, particularly given that researchers often do not know anything about the true frequency of the data generating process. We leave the process of determining a method to characterize the true frequency of a process to future work.

## References

- Ahmad, Y. and W. Craighead, 2011. Temporal Aggregation and Purchasing Power Parity Persistence. *Journal of International Money and Finance*, 30, 817-830.
- Amemiya, .T. and Wu, R., 1972. The Effect of Aggregation on Prediction in the Autoregressive Model. *Journal of the American Statistical Association*, 67 (339), 628 - 632.
- Belaire-Franch, J., Opong, K.K., 2005. Some evidence of random walk behavior of Euro exchange rates using ranks and signs. *Journal of Banking and Finance* 29, 1631-1643.
- Brewer, K. 1973. Some Consequences of Temporal Aggregation and Systematic Sampling for ARMA and ARMAX Models. *Journal of Econometrics*, 1, 133 - 154.
- Campbell, J. Y. and N. G. Mankiw, 1987. Are Output Fluctuations Transitory? *Quarterly Journal of Economics*, 102 (4), 857 - 880.
- Campbell, J. Y. and N. G. Mankiw, 1989. International Evidence on the Persistence of Economic Fluctuations. *Journal of Monetary Economics*, 23, 319 - 333.
- Campos, J., Ericsson, N., and Hendry, D., 1990. An Analogue Model of Phase-Averaging Procedures. *Journal of Econometrics*, 43, 275 - 292.
- Cecchetti, S.G. and P. Lam, 1994. Variance-ratio Tests: Small-sample Properties with an Application to International Output Data. *Journal of Business and Economic Statistics*, 12, 177-186.
- Choi, C., N. Mark and D. Sul, 2006. Unbiased Estimation of the Half-Life to PPP Convergence in Panel Data. *Journal of Money Credit and Banking*, 38, 921-938.
- Chow, K.V., K.C. Denning, 1993. A simple multiple variance ratio test. *Journal of Econometrics* 58, 385-401.
- Christiano, L. J., M. Eichenbaum and D. Marshall, 1991. The Permanent Income Hypothesis Revisited. *Econometrica*, 59, 397-423.
- Cochrane, J. 1988. How big is the random walk in GNP?. *Journal of Political Economy*, 96,

893-920.

Cochrane, J. 2001. *Asset Pricing*. Princeton University Press: Princeton.

Deaton, A., 1987. Life-cycle Models of Consumption: Is the Evidence Consistent with the Theory?, in *Advances in Econometrics - Fifth World Congress*, Vol. II, ed T F Bewley, Cambridge, U.K., Cambridge University Press, 121 - 148.

Fama, E.F., 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25, 383-417.

Fong, W.M., Koh, S.K., Ouliaris, S., 1997. Joint variance-ratio tests of the martingale hypothesis for exchange rates. *Journal of Business & Economic Statistics* 15 (1), 51–59.

Grilli, V. and Kaminsky, G., 1991. Nominal exchange rate regimes and the real exchange rate. *Journal of Monetary Economics*, 27, 191 - 212.

Haug, A., 1991. The Random Walk Hypothesis of Consumption and Time Aggregation. *Journal of Macroeconomics*, 13 (4), 691-700.

Huizinga, John, 1987, An empirical investigation of the long-run behavior of real exchange rates, *Carnegie-Rochester Conference Series on Public Policy* 27, 149-214.

Lettau, M., and Ludvigson, S., 2010. Measuring and Modeling Variation in the Risk-Return Trade-off. *Handbook of Financial Econometrics*, vol. 1, 617–690, ed. by Yacine Ait-Shalia and Lars-Peter Hansen. Elsevier.

Liu, C.Y., and He, J., 1991. A variance ratio test of random walks in foreign exchange rates. *Journal of Finance* 46, 773–785.

Lo, A.W., 2002. The Statistics of Sharpe ratios, *Financial Analysts Journal* 58 (4), 36-52

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

Lo, A.W., MacKinlay, A.C., 1989. The size and power variance ratio test in finite samples: A

- Monte Carlo investigation. *Journal of Econometrics* 40, 203–238.
- Nelson, C. and Plosser, C., 1982. Trends and random walks in macroeconomic time series: Some evidence and implications. *Journal of Monetary Economics*, 10 (2), 139 - 162.
- Paya, I., A. Duarte and K. Holden, 2007. On the Relationship between Inflation Persistence and Temporal Aggregation. *Journal of Money Credit and Banking*, 39, 1521-1531.
- Paya, I., and D. Peel, 2006. Temporal Aggregation of an ESTAR process: some implications for purchasing power parity adjustment. *Journal of Applied Econometrics*, 21 (5), 655 - 668.
- Rossana, R. J. and J. J. Seater, 1992. Aggregation, Unit Roots, and the Time-Series Structure of Manufacturing Real Wages. *International Economic Review*, 33, 159-179.
- Rossana, R. J. and J. J. Seater, 1995. Temporal Aggregation and Economic Time Series. *Journal of Business and Economic Statistics*, 13, 441-451.
- Sargan, J. D., 1974. Some Discrete Approximations to Continuous Time Stochastic Models. *Journal of the Royal Statistical Society, Series B (Methodological)*, 36 (1), 74 - 90.
- Siegel, J. J., 1998. *Stocks for the Long Run: The Definitive Guide to Financial Market Returns and Long-term Investment Strategies*, Seconded. NewYork, NY: McGraw Hill.
- Silvestrini, A. and D. Veredas, 2008. Temporal Aggregation of Univariate and Multivariate Time Series Models: A Survey. *Journal of Economic Surveys*, 22 (3), 458 - 497.
- Taylor, A., 2001. Potential Pitfalls for the Purchasing Power Parity Puzzle? Sampling and Specification Biases in Mean Reversion Tests of the Law of One Price. *Econometrica*, 69, 473-498.
- Telser, L. 1967. Discrete Samples and Moving Sums in Stationary Stochastic Processes. *Journal of the American Statistical Association*, 62, 484 - 499.
- Tiao, G. C., 1972. Asymptotic behavior of temporal aggregates of time series. *Biometrika*, 59, 525 - 531.
- Working, H., 1960. Note on the Correlation of First Differences of Averages in a Random Chain.

Econometrica, 28, 916-918.

Yilmaz, K., 2003. Martingale property of exchange rates and central bank interventions. *Journal of Business & Economic Statistics* 21 (3), 383-395.

Zellner, A., 1966. On the analysis of first order autoregressive models with incomplete data. *International Economic Review*, 7, 72 - 76.

Table 1. *VR* tests for the dollar/sterling spot exchange rate

Sample: August 7, 1974 to May 13, 2011.

	Time Average			Interval Sampling	
	Daily	Weekly	Monthly	Weekly	Monthly
Panel A. Variance Ratio					
Chow and Denning	3.42*	6.59*	5.47*	1.77	1.96
Lo and MacKinlay					
Lag $k = 2$	1.05*	1.23*	1.36*	1.01	1.10
Lag $k = 4$	1.07*	1.36*	1.62*	1.03	1.21
Lag $k = 8$	1.08*	1.53*	1.79*	1.13	1.27
Lag $k = 16$	1.09*	1.71*	1.86*	1.24	1.29
Panel B. Autocorrelation Coefficient of currency returns					
Lag $q = 1$	0.047	0.227	0.355	0.009	0.094
Lag $q = 2$	0.002	-0.003	0.034	-0.009	0.039
Lag $q = 3$	-0.013	0.046	0.067	0.051	0.037
Lag $q = 4$	0.007	0.041	0.042	0.038	-0.001
Lag $q = 8$	-0.003	-0.000	-0.001	-0.032	0.010

Notes: Numbers in table are variance ratios. An asterisk denotes rejection of the null

of the heteroskedasticity robust Chow and Denning (1993) and

Lo and MacKinlay (1988) statistics at the 5% level for lags 2, 4, 8, 16

Table 2. Dollar/sterling spot August 7, 1974 to May 13, 2011.

Autocorrelation Coefficient of order  $q$  for the  $k$ -differenced series

$k, q$	RW	Time Average			Interval Sampling	
		Daily	Weekly	Monthly	Weekly	Monthly
$k = 2 \ q = 1$	0.500	0.523	0.591	0.644	0.499	0.560
$k = 2 \ q = 2$	0	0.018	0.109	0.181	0.021	0.095
$k = 2 \ q = 4$	0	0.003	0.057	0.045	0.058	0.021
$k = 4 \ q = 1$	0.750	0.767	0.824	0.850	0.766	0.790
$k = 4 \ q = 2$	0.500	0.510	0.575	0.595	0.539	0.552
$k = 4 \ q = 4$	0	0.014	0.124	0.088	0.094	0.030

Notes: Numbers in table are correlation coefficients of order  $q$  for the  $k$ -differenced series. The second column reports the values that would correspond to a Random Walk model of the series in levels. The third column are the values for the daily exchange rate and subsequent columns for the time average or interval sampled of the daily figures.

Table 3: Impact on Adjusted Sharpe Ratios from temporal aggregation

**Panel A: Descriptive Stats**

	Mean	Variance
Actual (Daily) Data	0.00025978	0.00011323

**Panel B: Modified Sharpe Ratio for Time Averaged Data**

Holding Period Horizon (years)	Temporal Aggregation ( $h$ )			
	( $h = 1$ )	( $h = 20$ )	( $h = 60$ )	( $h = 240$ )
	Daily	Monthly	Quarterly	Annual
1	0.0248	0.1110	0.1992	0.3547
3	0.0220	0.0911	0.1670	0.3441
5	0.0232	0.0801	0.1474	0.3132
10	0.0205	0.0741	0.1376	0.2779

Notes: Daily returns were constructed from the closing price of the S&P 500 based on daily price data

between 1970 - 2013. Temporal aggregation was based on values of  $h = \{1, 20, 60, 240\}$ . The

numbers in Panels B equal the modified Sharpe ratio:  $\zeta^* \equiv \frac{S_k^*}{\sqrt{k}}$ .



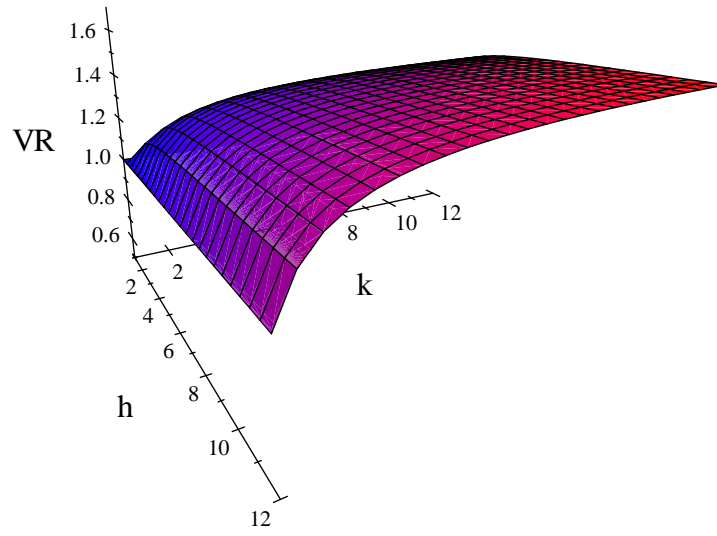


Figure 1: Plot of the  $VR$  for different values of  $h$  and  $\tilde{k}$ .

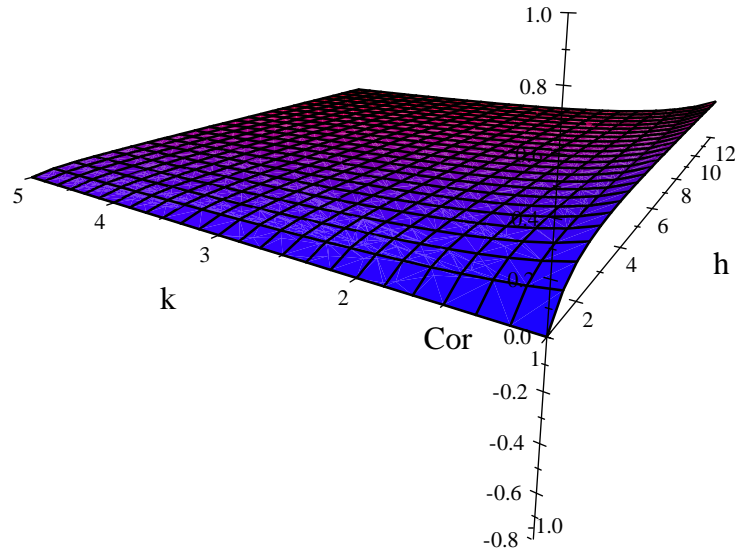


Figure 2: Correlation coefficient in the case of  $\tilde{k} = \tilde{q}$

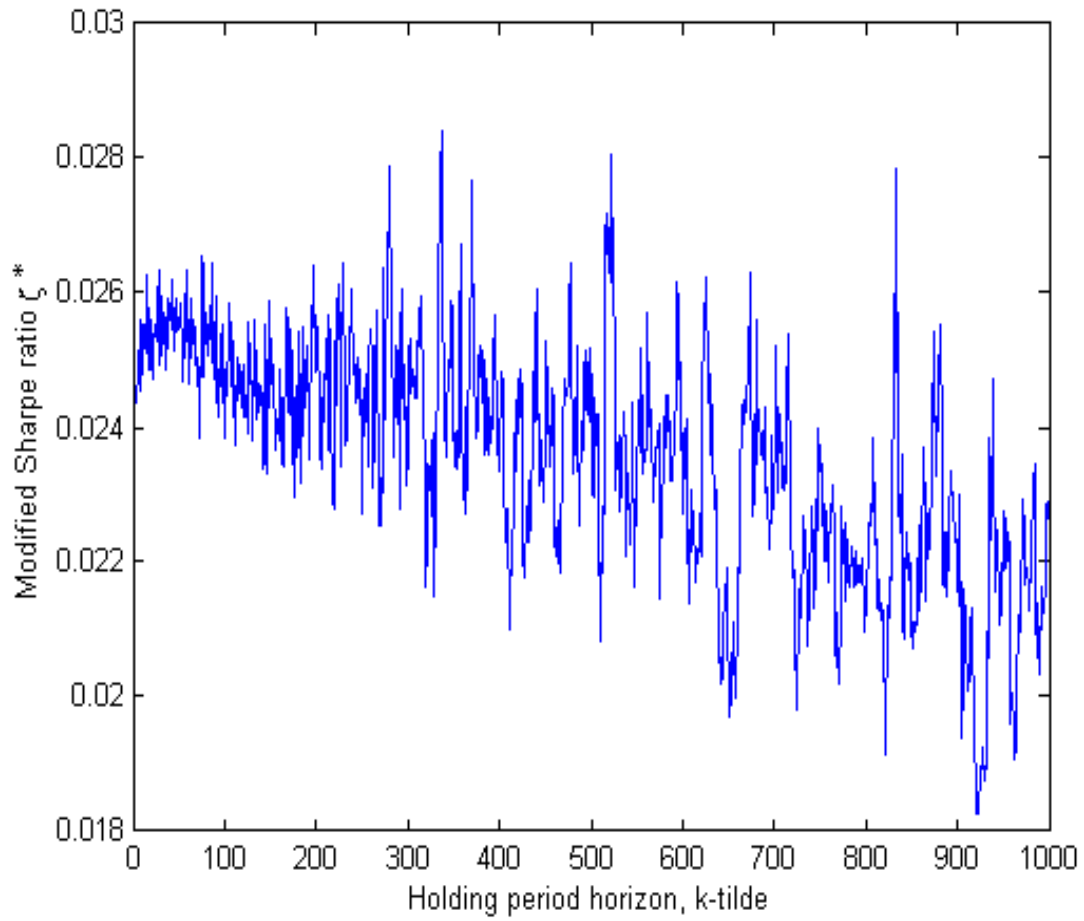


Figure 3: Value of the modified Sharpe ratio,  $\zeta^*$  for different values of  $\tilde{k}$  for daily net returns data.

# Temporal Aggregation of Random Walk Processes and Implications for Asset Prices

## Technical Appendix

Yamin S Ahmad\*

University of Wisconsin - Whitewater

Ivan Paya<sup>†</sup>

Lancaster University Management School, Economics Department

### Abstract

This paper is the technical appendix for the paper “Temporal Aggregation of Random Walk Processes and Implications for Asset Prices”. It contains details and derivations of some of the analytical expressions in the paper, which the referees may find useful in reading the manuscript. It is not intended for publication.

Current version: March 7th, 2014

---

\* Department of Economics, 800 W Main Street, Whitewater, WI 53190 Email: [ahmady@uww.edu](mailto:ahmady@uww.edu), Homepage: <http://facstaff.uww.edu/ahmady/> Tel: (262) 472 5576, Fax: (262) 472 4683

<sup>†</sup> Lancaster University Management School, Lancaster, LA1 4YX, UK Email: [i.paya@lancaster.ac.uk](mailto:i.paya@lancaster.ac.uk)

Consider a series  $y_t$  that follows a random walk, where increments are uncorrelated:

$$y_t = y_{t-1} + \varepsilon_t = y_0 + \sum_{s=1}^t \varepsilon_s$$

where  $\varepsilon_t$  is white noise:  $E(\varepsilon_t) = 0$ ;  $Var(\varepsilon_t) = E(\varepsilon_t^2) = \sigma_\varepsilon^2$ ;  $E(\varepsilon_t \varepsilon_s) = 0 \forall t \neq s$  and without any loss of generality we assume the initial condition to be zero,  $y_0 = 0$ . Then the process that  $y_t$  follows has the following properties:

First order moments:  $E(y_t) = E[y_0 + \sum_{s=1}^t \varepsilon_s] = E(y_0) = y_0$ ;

Second order moments:

$$\begin{aligned} Var(y_t) &= E[y_t^2] - E[y_t]^2 = E\left[\sum_{s=1}^t \varepsilon_s^2\right] = \sum_{s=1}^t E(\varepsilon_s^2) = t\sigma_\varepsilon^2 \\ Cov(y_t, y_{t-k}) &= E\left[\left(y_0 + \sum_{s=1}^t \varepsilon_s\right)\left(y_0 + \sum_{s=1}^{t-k} \varepsilon_s\right)\right] - y_0^2 = (t-k)\sigma_\varepsilon^2 \\ Cor(y_t, y_{t-k}) &= \frac{(t-k)\sigma_\varepsilon^2}{(t\sigma_\varepsilon^2)^{0.5}((t-k)\sigma_\varepsilon^2)^{0.5}} = \frac{t-k}{t^{0.5}(t-k)^{0.5}} = \sqrt{\frac{t-k}{t}} \end{aligned}$$

Third order moments: Here we assume that the white noise process is a Gaussian white noise process, which means that  $E(\varepsilon_t^3) = 0$ ,  $E(\varepsilon_t^a \varepsilon_s^b) = 0$ .

$$\begin{aligned} E\left[(y_t - E(y_t))^3\right] &= E\left[\left(\sum_{s=1}^t \varepsilon_s\right)^3\right] \\ &= E\left[\sum_{s_1+s_2+\dots+s_t=n} \binom{n}{s_1, s_2, \dots, s_t} \prod_{1 \leq j \leq t} \varepsilon_j^{s_j}\right] \text{ by the multinomial theorem} \\ &= E\left[\left(\sum_{i=1}^t \varepsilon_i^3\right) + 3\left(\sum_{i=1}^t \prod_j \varepsilon_i^2 \varepsilon_j\right) + 6 \prod_{u,v,w} \varepsilon_u \varepsilon_v \varepsilon_w\right] \end{aligned}$$

This gives us the following measure of Skewness:

$$\gamma \equiv \frac{E\left[(y_t - E(y_t))^3\right]}{\left(E\left[(y_t - E(y_t))^2\right]\right)^{\frac{3}{2}}} = \frac{E\left[\left(\sum_{s=1}^t \varepsilon_s\right)^3\right]}{(t\sigma_\varepsilon^2)^{\frac{3}{2}}}$$

Fourth order moments:

Kurtosis:

$$\beta \equiv \frac{E \left[ (y_t - E(y_t))^4 \right]}{\left( E \left[ (y_t - E(y_t))^2 \right] \right)^2} = \frac{E \left[ \left( \sum_{s=1}^t \varepsilon_s \right)^4 \right]}{(t\sigma_\varepsilon^2)^2}$$

### Time Averaging Case

Let  $\tilde{t}$  represent observations from the aggregated data, where  $\tilde{t} = \frac{t}{h}$ , and  $\tilde{k}$  represent lags of the aggregated series, i.e.  $\tilde{k} = \frac{k}{h}$ . Then the aggregated series can be represented in terms of the original observations as  $y_{\tilde{t}}^* = \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-i}$ . Similarly, the  $\tilde{k}'th$  lag is  $y_{\tilde{t}-\tilde{k}}^* = \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-h\tilde{k}-i} = \frac{1}{h} \sum_{i=0}^{h-1} y_{h(\tilde{t}-\tilde{k})-i}$ .

This aggregated series will have the following properties:  $E(y_{\tilde{t}}^*) = \frac{1}{h} E \left[ \sum_{i=0}^{h-1} \left( y_0 + \sum_{s=1}^{h\tilde{t}-i} \varepsilon_s \right) \right] = y_0$ ;

Variance:

$$\begin{aligned} Var(y_{\tilde{t}}^*) &= Var \left( \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-i} \right) = \frac{1}{h^2} Var \left( \sum_{i=0}^{h-1} \sum_{s=1}^{h\tilde{t}-i} \varepsilon_s \right) \\ &= \frac{1}{h^2} Var \left[ h \sum_{j=1}^{t-(h-1)} \varepsilon_j + \sum_{i=1}^{h-1} i \varepsilon_{t+1-i} \right] \\ &= \sigma_\varepsilon^2 \left[ (t - (h - 1)) + \frac{(h - 1)(2h - 1)}{6h} \right] = \sigma_\varepsilon^2 \left[ (h\tilde{t} - (h - 1)) + \frac{(h - 1)(2h - 1)}{6h} \right] \end{aligned}$$

Covariance:

$$\begin{aligned} Cov(y_{\tilde{t}}^*, y_{\tilde{t}-\tilde{k}}^*) &= Cov \left[ \left( \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-i} \right) \left( \frac{1}{h} \sum_{i=0}^{h-1} y_{h(\tilde{t}-\tilde{k})-i} \right) \right] = Cov \left[ \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=1}^{h\tilde{t}-i} \varepsilon_s, \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=1}^{h(\tilde{t}-\tilde{k})-i} \varepsilon_s \right] \\ &= \sigma_\varepsilon^2 \left[ t - k - (h - 1) + \frac{(h - 1)(2h - 1)}{6h} \right] = \sigma_\varepsilon^2 \left[ h(\tilde{t} - \tilde{k}) - (h - 1) + \frac{(h - 1)(2h - 1)}{6h} \right] \end{aligned}$$

Correlation:

$$\begin{aligned}
Cor\left(y_t^*, y_{t-\tilde{k}}^*\right) &= \frac{Cov\left(y_t^*, y_{t-\tilde{k}}^*\right)}{\left(Var\left(y_t^*\right)\right)^{0.5}\left(Var\left(y_{t-\tilde{k}}^*\right)\right)^{0.5}} \\
&= \frac{\sigma_\varepsilon^2\left[h\left(\tilde{t}-\tilde{k}\right)-\left(h-1\right)+\frac{\left(h-1\right)\left(2h-1\right)}{6h}\right]}{\sqrt{\sigma_\varepsilon^2\left[\left(h\tilde{t}-\left(h-1\right)\right)+\frac{\left(h-1\right)\left(2h-1\right)}{6h}\right]}\cdot\sqrt{\sigma_\varepsilon^2\left[\left(h\left(\tilde{t}-\tilde{k}\right)-\left(h-1\right)\right)+\frac{\left(h-1\right)\left(2h-1\right)}{6h}\right]}} \\
\Rightarrow \rho_{\tilde{k}}^* &= \sqrt{\frac{h\left(\tilde{t}-\tilde{k}\right)+D}{h\tilde{t}+D}}
\end{aligned}$$

where  $D(h) = \frac{(h-1)(2h-1)}{6h} - (h-1) = -\frac{(h-1)}{6h} [1 + 4h] < 0$  since  $h \geq 1$ .

### Interval Sampling Case

The aggregated series can be represented in terms of the original observations as  $y_t^* = y_{h\tilde{t}}$ . Similarly, the  $\tilde{k}'th$  lag of  $y_t^*$  is  $y_{t-\tilde{k}}^* = y_{h\tilde{t}-h\tilde{k}} = y_{h(\tilde{t}-\tilde{k})}$ . This aggregated series will have the following properties:  $E\left(y_t^*\right) = y_0 + \sum_{s=1}^{\tilde{t}} E\left(\varepsilon_{hs}\right) = y_0$ ;  $Var\left(y_t^*\right) = Var\left(\sum_{s=1}^{\tilde{t}h} \varepsilon_s\right) = \tilde{t}h\sigma_\varepsilon^2$ ;  $Cov\left(y_t^*, y_{t-\tilde{k}}^*\right) = Cov\left[\sum_{s=1}^{\tilde{t}h} \varepsilon_s, \sum_{s=1}^{(\tilde{t}-\tilde{k})h} \varepsilon_s\right] = \left(\tilde{t}-\tilde{k}\right)h\sigma_\varepsilon^2$ ;  $Cor\left(y_t^*, y_{t-\tilde{k}}^*\right) = \rho_{\tilde{k}}^* = \frac{h\left(\tilde{t}-\tilde{k}\right)\sigma_\varepsilon^2}{\sqrt{h\tilde{t}\sigma_\varepsilon^2}\sqrt{h\left(\tilde{t}-\tilde{k}\right)\sigma_\varepsilon^2}} = \sqrt{\frac{\left(\tilde{t}-\tilde{k}\right)}{\tilde{t}}}$

### Variance Ratio Test

The true DGP we are considering exhibits uncorrelated increments  $\Delta y_t = \varepsilon_t$ , where  $\Delta$  is the first difference operator, and  $Var(\Delta y_t) = \sigma_\varepsilon^2$ . The  $k$ -difference of  $y_t$ ,  $\Delta_k y_t = \sum_{i=0}^{k-1} \varepsilon_{t-i}$ , and therefore  $Var(\Delta_k y_t) = k\sigma_\varepsilon^2$ .

The time average aggregated series can be represented in terms of the original observations (assuming without loss of generality that initial condition  $y_0 = 0$ ):

$$y_t^* = \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-i} = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=1}^{h\tilde{t}-i} \varepsilon_s$$

Similarly, the  $\tilde{k}'th$  lag of  $y_t^*$  is:

$$y_{t-\tilde{k}}^* = \frac{1}{h} \sum_{i=0}^{h-1} y_{h\tilde{t}-h\tilde{k}-i} = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=1}^{h(\tilde{t}-\tilde{k})-i} \varepsilon_s$$

The  $\tilde{k}$ -difference of this process is equal to

$$\begin{aligned}\Delta_{\tilde{k}}y_t^* &= y_t^* - y_{t-\tilde{k}}^* = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=h(\tilde{t}-\tilde{k})+1-i}^{h\tilde{t}-i} \varepsilon_s \\ &= \frac{1}{h} \left[ h \sum_{j=t-k+1}^{t-(h-1)} \varepsilon_j + \sum_{i=1}^{h-1} i \varepsilon_{t-i+1} + \sum_{i=1}^{h-1} i \varepsilon_{t-k-i+1} \right]\end{aligned}$$

and its variance is

$$\begin{aligned}\text{Var}(\Delta_{\tilde{k}}y_t^*) &= \frac{1}{h^2} \left( h^2 \sum_{j=t-k+1}^{t-(h-1)} \sigma_\varepsilon^2 + \sigma_\varepsilon^2 \sum_{i=1}^{h-1} i^2 + \sigma_\varepsilon^2 \sum_{i=1}^{h-1} i^2 \right) \\ &= \frac{1}{h^2} \left( h^2 \sum_{j=t-k+1}^{t-(h-1)} \sigma_\varepsilon^2 + 2\sigma_\varepsilon^2 \sum_{i=1}^{h-1} i^2 \right) \\ &= \left( (k-h+1) + \frac{(2h-1)(h-1)}{3h} \right) \sigma_\varepsilon^2 \\ &= \left( h(\tilde{k}-1) + 1 + \frac{(2h-1)(h-1)}{3h} \right) \sigma_\varepsilon^2 \\ &= \left( h\tilde{k} + 2D + (h-1) \right) \sigma_\varepsilon^2\end{aligned}$$

And we obtain the expression for the Variance Ratio

$$\text{VR}(\tilde{k}) = \frac{\text{Var}(\Delta_{\tilde{k}}y_t^*)}{\tilde{k}\text{Var}(\Delta y_t^*)} = \frac{h(\tilde{k}-1) + 1 + \frac{(2h-1)(h-1)}{3h}}{\tilde{k} \left( 1 + \frac{(2h-1)(h-1)}{3h} \right)} = \frac{h\tilde{k} + 2D + (h-1)}{\tilde{k} (2D + (2h-1))}$$

### Autocorrelations of Differenced Series:

We introduce the parameter  $\tilde{q}$  which is the order of the autocovariance for the  $\tilde{k}$ -differenced aggregated series. Remember that  $\Delta y_t^* = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=t-(h-1)-i}^{t-i} \varepsilon_s$ , and, similarly,  $\Delta y_{t-\tilde{q}}^* = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=h(\tilde{t}-\tilde{q})-h+1-i}^{h(\tilde{t}-\tilde{q})-i} \varepsilon_s = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=t-h\tilde{q}-(h-1)-i}^{t-h\tilde{q}-i} \varepsilon_s$ . Given that  $\Delta y_t^*, \Delta y_{t-\tilde{q}}^*$  are only sums of white noise processes then  $\text{Cov}(\Delta y_t^*, \Delta y_{t-\tilde{q}}^*) = E \left[ \Delta y_t^* \Delta y_{t-\tilde{q}}^* \right]$ .

For the product to have terms in comon the following needs to hold:  $2(h-1) \geq h\tilde{q}$ , or,  $h(2-\tilde{q}) \geq 2$ . This only holds for the case of  $\tilde{q} = 1$  given that  $h \geq 1$ . So the  $\text{Cov}(\Delta y_t^*, \Delta y_{t-\tilde{q}}^*)$  only has solution different than zero for  $\tilde{q} = 1$ . This result coincides with Working (1960).

$$\begin{aligned}
Cov(\Delta y_t^*, \Delta y_{t-\tilde{q}}^*) &= \frac{1}{h^2} \left[ \sum_{i=1}^{h-1} i(h-i) \sigma_\varepsilon^2 \right] = \frac{1}{h^2} \left[ \frac{h(h-1)(h-2)}{6} \sigma_\varepsilon^2 \right] \\
&= \frac{(h-1)(h+1)}{6h} \sigma_\varepsilon^2 = \frac{(h^2-1)}{6h} \sigma_\varepsilon^2
\end{aligned}$$

Let us now compute  $Cov(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*)$ . Note first that we can write

$$\begin{aligned}
\Delta_{\tilde{k}} y_t^* &= \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=h(\tilde{t}-\tilde{k})+1-i}^{\tilde{h}\tilde{t}-i} \varepsilon_s = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=h\tilde{t}-h\tilde{k}+1-i}^{\tilde{h}\tilde{t}-i} \varepsilon_s \\
&= \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=t-k+1-i}^{t-i} \varepsilon_s
\end{aligned}$$

and similarly

$$\begin{aligned}
\Delta_{\tilde{k}} y_{t-\tilde{q}}^* &= \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=h(\tilde{t}-\tilde{q}-\tilde{k})+1-i}^{h(\tilde{t}-\tilde{q})-i} \varepsilon_s = \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=h(\tilde{t}-\tilde{q})-h\tilde{k}+1-i}^{h\tilde{t}-h\tilde{q}-i} \varepsilon_s \\
&= \frac{1}{h} \sum_{i=0}^{h-1} \sum_{s=t-k+1-i-h\tilde{q}}^{t-i-h\tilde{q}} \varepsilon_s
\end{aligned}$$

The covariance is derived from the product of  $\Delta_{\tilde{k}} y_t^*$  and  $\Delta_{\tilde{k}} y_{t-\tilde{q}}^*$ . The condition to have terms in common is  $-k+1-(h-1) \leq -h\tilde{q}$ , or  $h(1+\tilde{k}-\tilde{q}) \geq 2$ . This condition holds when  $\tilde{k} \geq \tilde{q}$  given that  $h > 1$ .

The covariance now has two expressions depending on the value of  $\tilde{k}$  and  $\tilde{q}$  (remember that  $\tilde{k} \geq \tilde{q}$  has to hold):

$$Cov(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*) = \frac{1}{h^2} \left[ \sum_{i=1}^{h-1} i(h-i) \sigma_\varepsilon^2 \right] = \frac{(h^2-1)}{6h} \sigma_\varepsilon^2 \text{ if } \tilde{q} = \tilde{k}$$

So for any value of  $\tilde{k}$ , and  $\tilde{q}$  the distortion calculated in Working holds as long as  $\tilde{q} = \tilde{k}$ .

The second expression



$$\begin{aligned}
Cov(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*) &= \frac{1}{h^2} \left[ \left[ h(\tilde{k} - \tilde{q} - 1) + 1 \right] h^2 + 2 \sum_{i=1}^{h-1} ih \right] \sigma_\varepsilon^2 \\
&= \frac{1}{h^2} \left[ \left[ h(\tilde{k} - \tilde{q} - 1) + 1 \right] h^2 + h^2(h-1) \right] \sigma_\varepsilon^2 \\
&= \left[ h(\tilde{k} - \tilde{q} - 1) + 1 + (h-1) \right] \sigma_\varepsilon^2 = h(\tilde{k} - \tilde{q}) \sigma_\varepsilon^2 \text{ if } \tilde{k} > \tilde{q}
\end{aligned}$$

Correlation. For the case of  $\tilde{k} = 1$  (and therefore  $\tilde{q} = 1$ )

$$Cor(\Delta y_t^*, \Delta y_{t-\tilde{q}}^*) = \frac{\frac{h^2-1}{6h}}{\frac{2h^2+1}{3h}} = \frac{\frac{h^2-1}{6h}}{1 + \frac{(2h-1)(h-1)}{3h}} = \frac{h^2-1}{2(2h^2+1)} = \frac{h^2-1}{4h^2+2}$$

For the case of  $\tilde{k} > 1$  we have two cases

$$Cor(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*) = \frac{\frac{h^2-1}{6h}}{h(\tilde{k}-1) + 1 + \frac{(2h-1)(h-1)}{3h}} \text{ if } \tilde{q} = \tilde{k}$$

For the case when  $\tilde{k} > \tilde{q}$

$$Cor(\Delta_{\tilde{k}} y_t^*, \Delta_{\tilde{k}} y_{t-\tilde{q}}^*) = \frac{h(\tilde{k} - \tilde{q})}{h(\tilde{k} - 1) + 1 + \frac{(2h-1)(h-1)}{3h}} = \frac{h(\tilde{k} - \tilde{q})}{h\tilde{k} - (h-1) + \frac{(2h-1)(h-1)}{3h}} \text{ if } \tilde{k} > \tilde{q}$$