Persistence under Temporal Aggregation and Differencing

Uwe Hassler[*]
Goethe University Frankfurt[†]

August 8, 2013

Abstract

Temporal aggregation is known to affect the persistence of time series, where persistence reflects, roughly speaking, the degree of positive serial correlation. It is suggested to measure persistence as percentage of the long-run variance not attributable to the variance ("long-run variance difference ratio"). Here we study the aggregation of flow variables as well as stock data, and difference-stationarity is allowed for. Moreover, moving averages encountered e.g. when computing annual growth rates (seasonal differences) are investigated. It is clarified when persistence is increased or decreased, and by how much. Our results are exact for finite aggregation level. They are illustrated with numerical examples and real data. Approximate results for growing aggregation level are provided, too.

[*]This research was carried out within the project HA 3306/5-1 financed by the German Research Foundation (DFG). Earlier versions of this paper were written while visiting the Department of Mathematics and Statistics at University of Sydney, the Department of Econometrics and Business Statistics at Monash University, Melbourne, and the Department of Finance and Accounting at the Vietnamese-German University, Ho-Chi-Minh City, and they were presented at seminars at Hitotsubashi University, Tokyo, Sogang University, Seoul, and Freie Universität Berlin. I thank In Choi, Myriam Hadnes, Dieter Nautz, Shelton Peiris, Henghsiu Tsai, Farshid Vahid, and Mu-Chun Wang for support and many helpful comments.

[†]RuW Building, Grueneburgplatz 1, 60323 Frankfurt, Germany, tel: +49-69-798-34762. Email: hassler@wiwi.uni-frankfurt.de

Keywords: cumulating, skip sampling, difference-stationarity, seasonal differences

JEL classification: C22 (Time-Series Models), C82 (Methodology for Collecting, Estimating, and Organizing Macroeconomic Data)

1 Introduction

Temporal aggregation is a standard routine when working with time series. Even if the researcher does not aggregate him- or herself, the statistical offices making available the data may only provide aggregates. Many stationary time series display positive autocorrelation such that subsequent observations form clusters: positive observations tend to be followed by positive ones, while negative observations tend to induce negative ones. By persistence we understand roughly speaking the strength of such a tendency, which depends not only on the autocorrelation coefficient at lag one but also on higher order lags.

It has been empirically documented e.g. by Rossana and Seater (1995) that temporal aggregation affects persistence measures. But there is no clear answer whether persistence is increased or decreased, and by how much. In this paper we give an answer to these questions for the ratio of the difference of the (long-run) variance as a relative persistence measure. The results are exact for finite aggregation level and can be approximated with growing level. They are illustrated with numerical examples, and it is demonstrated with real data that they may explain aggregational effects in practice. Therefore, we also allow for nonstationary series where differencing is required to obtain stationarity. Moreover, we study the situation where seasonal growth rates are stationary, but annual growth rates (seasonal differences) are computed for convenience. Hassler and Demetrescu (2005) demonstrated experimentally and empirically that annual growth rates will dramatically exaggerate the degree of persistence relative to the persistence present in the seasonal rates. Their arguments are reinforced here theoretically in a general framework.

We study the aggregation of both, stock variables and flow variables. Typical flow data are monthly consumption, where temporal aggregation to quarterly or annual data consists of cumulating the monthly flows to the total quarterly or annual flow. Typical stock series are daily prices or interest rates. In order to obtain e.g. weekly data, one may compute the average of all days of the week, or alternatively, one may take the last weekday as representative of the whole week (called skip sampling).

Econometricians started investigations into the effect of temporal aggregation on the dynamic structure of time series a long time ago. Early results for random walks were obtained by Working (1960), and for autoregressive moving-average (ARMA) models by Brewer (1973) and Weiss (1984). A general treatment of integrated (of order one) ARIMA models was provided by Wei (1981) and Stram and Wei (1986); see also the review chapter in Wei (1990). There seem to be no theoretical contributions, however, attacking the effect of temporal aggregation on persistence. An early exception is related research for growing aggregation level by Tiao (1972), who proves that the cumulation of a stationary series turns into white noise with growing aggregation level. This suggests that cumulation reduces persistence at least asymptotically with growing aggregation level. In summary of this literature, it has been claimed that the effect of temporal aggregation on persistence in finite samples is an empirical matter, see for instance Rossana and Seater (1995), and more recently Paya, Duarte and Holden (2007), who employed several widely used persistence measures as the sum of autoregressive coefficients, the largest autoregressive root and (cumulated) impulse responses: "statistical theory is not definite because some of the results are asymptotic and leave open the question what will happen with actual data", Rossana and Seater (1995, p. 443).

Here, we argue that exact theoretical results on the aggregational effect with respect to persistence are readily available when using *relative* measures, since aggregation affects both, the serial correlation and the variance. We study the long-run variance ratio, which technically equals the normalized spectrum at frequency zero and has been applied e.g. by Cogley and

Sargent (2005), see also Cochrane (1988). We find that persistence is not a characteristic property of an underlying process, but necessarily linked to the frequency of sampling. In particular, we learn for aggregation of (integrated) stock variables that skip sampling affects persistence very differently from averaging.

The next section provides the notation, the assumptions and a way to measure persistence, while Section 3 briefly reviews alternative aggregation schemes. The fourth section contains the theoretical results on aggregation and differencing with numerical illustrations. The theoretical results manage to explain the aggregational effects observed with real data in Section 5. The final section discusses implications for applied work in general. Proofs are relegated to the Appendix.

2 Persistence

2.1 Under stationarity

The univariate time series data are assumed to be generated by a (covariance) stationary process $\{y_t\}$, where the autocovariances at lag h are denoted as

$$\gamma(h) = \mathrm{E}[(y_t - \mathrm{E}(y_t))(y_{t+h} - \mathrm{E}(y_{t+h}))] = \gamma(-h)\,. \tag{1}$$

The expectation $\mu_t = \mathrm{E}(y_t)$ may be constant, or e.g. seasonally varying, or given by a time trend. The stochastic deviations are assumed to follow a regular linear process with an absolutely summable sequence of impulse responses $\{c_j\}$.

Assumption 1 *The process $\{y_t\}$, $t \in \mathbb{Z}$, is given by*

$$y_t = \mu_t + \sum_{j=0}^{\infty} c_j \varepsilon_{t-j} \quad \textit{with} \ \sum_{j=0}^{\infty} |c_j| < \infty\,,\ c_0 = 1\,,\ \textit{and} \ \sum_{j=0}^{\infty} c_j \neq 0\,,$$

where $\{\varepsilon_t\}$ is a zero mean white noise process with variance $Var(\varepsilon_t) = \sigma^2$.

Next to the variance, $\mathrm{Var}(y_t) = \gamma(0)$, we define the long-run variance $\omega^2 = \mathrm{LrV}(y_t)$ depending on all temporal correlation summarized in $\Gamma(y)$:

$$\mathrm{LrV}(y_t) = \sum_{h=-\infty}^{\infty} \gamma(h) = \mathrm{Var}(y_t) + \Gamma(y)\,, \quad \Gamma(y) = 2\sum_{h=1}^{\infty} \gamma(h)\,. \tag{2}$$

By Assumption 1, $\mathrm{LrV}(y_t)$ is positive and finite; the limiting cases of $\mathrm{LrV}(y_t) = 0$ or $\mathrm{LrV}(y_t) = \infty$ arising from fractional integration are covered in Souza (2005), Tsai and Chan (2005), and Hassler (2011). Under the above assumptions it holds that

$$\gamma(0) = \sigma^2 \sum_{j=0}^{\infty} c_j^2 \quad \text{and} \quad \mathrm{LrV}(y_t) = \sigma^2 \left(\sum_{j=0}^{\infty} c_j \right)^2 . \tag{3}$$

Campbell and Mankiw (1987) popularized the cumulated impulse responses as measure of persistence,

$$\mathrm{CIR}(y) = \sum_{j=0}^{\infty} c_j\,.$$

It does not rely on an autoregressive representation of finite order, and has further been advocated by Andrews and Chen (1994) as being superior to the largest autoregressive root in some cases. Note, however, that temporal aggregation will affect both, the variance and the autocovariances behind Γ, such that a priori the aggregational effect on $\mathrm{CIR}(y)$ is unclear.

In this paper, persistence is measured through the long-run variance difference ratio $\Pi = \mathrm{LVDR}$:

$$\Pi(y) := \mathrm{LVDR}(y) := \frac{\mathrm{LrV}(y_t) - \mathrm{Var}(y_t)}{\mathrm{LrV}(y_t)} = \frac{\Gamma(y)}{\mathrm{LrV}(y_t)} = 1 - \frac{\mathrm{Var}(y_t)}{\mathrm{LrV}(y_t)}\,. \tag{4}$$

A process is called *persistent* if $\Gamma(y) > 0$, and Π measures the percentage, the cumulated autocovariances $\Gamma(y)$ contribute to the long-run variance, or equivalently: the percentage of the long-run variance not attributable to the variance. If $\Gamma(y) < 0$, one may call the process *anti-persistent*, and $\Pi(y)$ measures by how much the variance exceeds the long-run variance, although the percentage interpretation is lost in this case. Closely related to the long-run variance difference ratio Π from (4) is the variance ratio by Cochrane (1988) containing essentially the same information

$$\mathrm{VR}(y) := \frac{\mathrm{LrV}(y_t)}{\mathrm{Var}(y_t)} = \frac{1}{1-\Pi(y)} \,. \tag{5}$$

Of course, $\mathrm{VR}(y)$ equals the normalized spectrum at frequency zero employed e.g. by Cogley and Sargent (2005).

Example 1 (ARMA(1,1)) Consider a stationary ARMA(1,1) process,

$$y_t = a\, y_{t-1} + \varepsilon_t + b\, \varepsilon_{t-1} = \varepsilon_t + (a+b) \sum_{j=1}^{\infty} a^{j-1} \varepsilon_{t-j} \,, \quad |a| < 1 \,,\ a \neq -b \,.$$

Clearly, for $a < 0$, the process tends to oscillate and hence shows no persistent behavior. In the case of $a > 0$, however, the process is all the more persistent the larger a is: y_t tends to stay close to the past level y_{t-1}, such that clusters of subsequent observations arise. The other side of the coin of persistence is that the effect of past shocks ε_{t-j} on the current level y_t dies out the more slowly the larger a is. The degree of persistence will be counter-balanced by negative values of b. We now quantify how Π reflects this feature:

$$\Pi(y) = \mathrm{LVDR}(y) = 1 - \frac{(1-a)\,(1+b^2+2ab)}{(1+a)\,(1+b)^2} \,. \tag{6}$$

In the white noise case ($a = b = 0$) the process is not persistent, $\Pi = 0$. But with increasing $0 < a$ the percentage of the long-run variance not attributable to the variance grows, up to $\Pi = 100\%$ in the limiting case of nonstationarity ($a = 1$). If the process is a pure invertible MA(1), $a = 0$, $|b| < 1$, the memory is much shorter and the impulse responses drop to zero after one period. Hence, strong persistence cannot be modelled by means of an MA(1) process. Although the long-run variance difference ratio is growing with b, $\Pi(y) = 2\,b(1+b)^{-2}$, it never exceeds 50%.

For applied work, Π has to be estimated from a sample of size T. The consistent estimation of a (long-run) variance from stationary data is a standard problem of course. Consistent long-run variance estimation is discussed e.g. in Hamilton (1994, Sect. 10.5).

2.2 Under nonstationarity

Clearly, many time series are not stationary. It is often assumed that the observed variable $\{z_t\}$ has to be differenced to obtain stationarity. With the usual difference operator Δ we define

$$\Delta^r z_t = y_t \,, \qquad t = 1, 2, \ldots, T \,, \tag{7}$$

for some natural number r, where $\{y_t\}$ is a (weakly) stationary sequence characterized in Assumption 1. In the case that $r = 0$, the observable $\{z_t\}$ itself is covariance stationary, while $r = 1$ gives $z_t - z_{t-1} = y_t$. In most applications the order of differencing is 1. In the case of nonstationarity ($r > 0$), the long-run variance difference ratio is computed in terms of stationary differences:

$$\Pi(\Delta^r z) = \mathrm{LVDR}(\Delta^r z) = 1 - \frac{\mathrm{Var}(\Delta^r z_t)}{\mathrm{LrV}(\Delta^r z_t)} \,.$$

In the case of $r = 1$, z_t may be decomposed into a random walk r_t and a cyclical component c_t integrated of order zero,

$$z_t = z_{t-1} + y_t = r_t + c_t \,,$$

although such a decomposition is not unique. Cochrane (1988), however, demonstrated that the input variance of the random walk is independently of the specific decomposition given as $\mathrm{Var}(\Delta r_t) = \mathrm{LrV}(\Delta z_t)$. Hence, $\Pi(\Delta z)$ measures the percentage of the random walk component exceeding the total input variance if $\mathrm{Var}(\Delta r_t) > \mathrm{Var}(\Delta z_t)$:

$$\Pi(\Delta z) = \frac{\mathrm{Var}(\Delta r_t) - \mathrm{Var}(\Delta z_t)}{\mathrm{Var}(\Delta r_t)} \,.$$

In many empirical applications, however, this ratio will be negative, simply meaning that the difference are not persistent: $\Gamma(\Delta z) < 0$.

3 Temporal aggregation and differencing

Let $\{z_t\}$, $t = 1, 2, \ldots, T$, denote a sample of univariate time series observations to be aggregated over m periods. We assume for simplicity that T is a multiple of m, $T = m\,N$. The aggregate is constructed for the new time scale τ. In the case of flow variables aggregation means cumulating m neighboring, non-overlapping observations to determine the total flow over m sub-periods,

$$\widetilde{z}_\tau \quad := \quad z_{m\tau} + z_{m\tau-1} + \ldots + z_{m(\tau-1)+1}\,, \quad \tau = 1, 2, \ldots, N\,, \tag{8}$$

where for the rest of the paper $m \geq 2$ is an integer. With stock data two aggregation schemes are encountered in practice. Often, stock variables are averaged, which is formally related to the cumulation of stocks with obvious notation: $\overline{z}_\tau = \widetilde{z}_\tau / m$. The usage of the new time scale τ indicates that the averages are not overlapping. Alternatively, stock variables are sometimes aggregated by systematic sampling or skip sampling where only every m'th data point is observed,

$$\dot{z}_\tau := z_{m\tau}, \quad \tau = 1, 2, \ldots, N\,. \tag{9}$$

If the basic variable $\{z_t\}$ is nonstationary as in (7), then the aggregates will be nonstationary, too. Let ∇ stand for the differences operating on the aggregate scale:

$$\nabla \widetilde{z}_\tau = \widetilde{z}_\tau - \widetilde{z}_{\tau-1} \quad \text{and} \quad \nabla \dot{z}_\tau = \dot{z}_\tau - \dot{z}_{\tau-1}\,.$$

For the differenced aggregates $\{\nabla^r \widetilde{z}_\tau\}$ and $\{\nabla^r \dot{z}_\tau\}$ we define the persistence measures as $\Pi(\nabla^r \widetilde{z})$ and $\Pi(\nabla^r \dot{z})$. Again, $r = 0$ refers to the situation where $z_t = y_t$ and hence $\widetilde{z}_\tau$ and $\dot{z}_\tau$ are stationary. Generally, it will make a difference whether one first aggregates and differences the aggregates, or the other way round, and the difference shall be spelled out in the next section.

Contrasting the case of non-overlapping averages, $\overline{z}_\tau$, we also consider moving averages of the following type:

$$S_m(L)\, z_t\,, \quad \text{where } S_m(L) = 1 + L + \cdots + L^{m-1}\,.$$

The polynomial $S_m(L)$ arises when computing seasonal differences $\Delta_m := 1 - L^m$ in order to obtain annual growth rates: $\Delta_m = S_m(L)\,\Delta$. Let us assume e.g. with $m = 12$ that the monthly growth rates $\Delta z_t = y_t$ are stationary. Economically or statistically, one is often interested in annual growth rates,

$$z_t - z_{t-12} = y_t + y_{t-1} + \cdots + y_{t-11}\,,$$

which are still stationary but considerably more persistent than the monthly rates y_t. The increase of persistence will be measured through $\Pi(\Delta_m z)$.

Now we are able to formalize the questions of this paper. Given an underlying process $\{z_t\}$ with persistence $\Pi(\Delta^r z)$, how do $\Pi(\nabla^r \widetilde{z})$ and $\Pi(\nabla^r \dot{z})$ computed from the r'th differences of the aggregates look like? What happens for m getting large? And, finally, by how much is the persistence increased when computed from annual growth rates rather than from seasonal rates?

4 Results

For the sake of clarity the presentation will be restricted to the empirically most relevant cases $r = 0$ or $r = 1$, although the proofs given in the Appendix treat the straightforward generalization for any integer $r \geq 0$.

We start with the case where stationary data are aggregated.

Proposition 1 *Let $z_t = y_t$, where $\{y_t\}$ satisfies Assumption 1 with autocovariances $\gamma(h)$. It then holds*

a) for $\{\dot{y}_\tau\}$:

$$\begin{aligned} \Pi(\dot{y}) &= 1 - \frac{m\; Var(y_t)}{2\pi \sum\limits_{j=0}^{m-1} f(2\pi j/m)} \qquad (10)\\ &\to 0 \quad \text{as } m \to \infty\,, \end{aligned}$$

where f is the spectrum defined as

$$f(\lambda) = \frac{1}{2\pi}\left(\gamma(0) + 2\sum_{h=1}^{\infty} \gamma(h)\cos(h\lambda)\right)\;;$$

b) for $\{\widetilde{y}_\tau\}$*:*

$$\begin{aligned}\Pi(\widetilde{y}) &= 1 - \frac{Var(y_t) + 2\sum_{h=1}^{m-1} \gamma(h)\,\frac{m-h}{m}}{LrV(y_t)} \qquad (11)\\ &\to 0 \quad as\ m \to \infty\,.\end{aligned}$$

Proof See Appendix.

The results hold irrespective of whether $\mathrm{LrV}(y_t) \geq \mathrm{Var}(y_t)$ or not. The interpretation below how persistence is affected by aggregation, however, relies on $\Pi \geq 0$.

Remark 1 From Proposition 1*a)* we observe that skip sampling stationary processes tends to reduce the persistence with growing m, which is of course very intuitive. For finite m, however, the exact effect is spelled out in (10). For an illustration with a stationary AR(1) process with $m = 1, 2, \ldots, 20$ see Figure 1, where $m = 1$ corresponds to no aggregation. With a moderate autoregressive parameter of $a = 0.7$ the percentage of the long-run variance not attributable to the variance drops from roughly 81% (for $m = 1$) to virtually 0 (for $m = 20$).

Remark 2 From Proposition 1*b)* we learn that cumulation of stationary processes tends to reduce the persistence with growing m, too, which is not surprising given that Tiao (1972) established that an MA(q) process turns into white noise with growing aggregation level, see also Hassler and Tsai (2013). For finite m we further quantify the exact effect in (11). It is illustrated with a stationary AR(1) process in Figure 2 for a moderate autoregressive parameter $a = 0.7$. The percentage of the long-run variance not attributable to the variance drops from roughly 81% (for $m = 1$) to 15% (for $m = 20$). So, the aggregational effect is not quite as strong as in the case of skip sampling.

Remark 3 Since we work with relative measures, the results with respect to cumulation continue to hold without modification in case of averaging stock variables:

$$\Pi(\nabla^r \widetilde{z}) = \Pi(\nabla^r \overline{z}) = 1 - \frac{\mathrm{Var}(\nabla^r \overline{z}_\tau)}{\mathrm{LrV}(\nabla^r \overline{z}_\tau)}\,,$$

where $r = 0$ in Proposition 1, or $r = 1$ below.

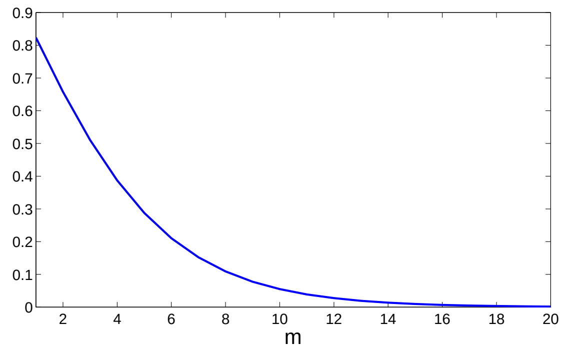

Figure 1. $\Pi(\dot{y})$ for $y_t = 0.7\, y_{t-1} + \varepsilon_t$

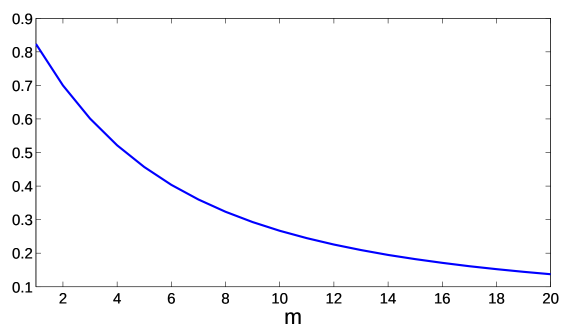

Figure 2. $\Pi(\widetilde{y})$ for $y_t = 0.7\, y_{t-1} + \varepsilon_t$

Proposition 2 *Let $\Delta z_t = y_t$, where $\{y_t\}$ satisfies Assumption 1 with autocovariances $\gamma(h)$. It then holds*

a) for $\{\nabla \dot{z}_\tau\}$:

$$\Pi(\nabla \dot{z}) = \Pi(\widetilde{y}) \,,$$

with $\Pi(\widetilde{y})$ from Proposition 1;

b) for $\{\nabla \widetilde{z}_\tau\}$:

$$\begin{aligned} \Pi(\nabla \widetilde{z}) &= 1 - \frac{Var(y_t) \sum_{i=0}^{2m-2} b_i^2 + 2 \sum_{h=1}^{2m-2} \left[\gamma(h) \sum_{i=0}^{2m-2-h} b_i b_{i+h} \right]}{m^3 \, LrV(y_t)} \qquad (12) \\ &\to \frac{1}{3} \quad as\ m \to \infty \,, \end{aligned}$$

where

$$b_0 = 1,\, b_1 = 2,\, \ldots, b_{m-1} = m,\, b_m = m-1, \ldots, b_{2m-2} = 1 \,.$$

Proof See Appendix.

Remark 4 Differencing and aggregation are not exchangeable. Think of an integrated process, $r = 1$ in (7). Differencing after cumulation corresponds to the case Proposition 2*b)*, while cumulating differences is covered by Proposition 1*b)*. Similarly, in case of skip sampling: $\Pi(\nabla \dot{z}) \neq \Pi(\dot{y})$.

Remark 5 With differenced aggregates the effect of cumulation depends, see Proposition 2*b)*: If $\{\Delta z_t\}$ has $\Pi(\Delta z) < 1/3$, then cumulation of the I(1) level will tend to exaggerate persistence (at least for large m), while for $\Pi(\Delta z) > 1/3$ a downward bias will be observed. Figure 3 depicts $\Pi(\nabla \widetilde{z})$ from (12) with a random walk: For $m = 1$, the differences are white noise, i.e. not persistent ($\Pi = 0$), but the limiting value 1/3 is approached very fast with growing m.

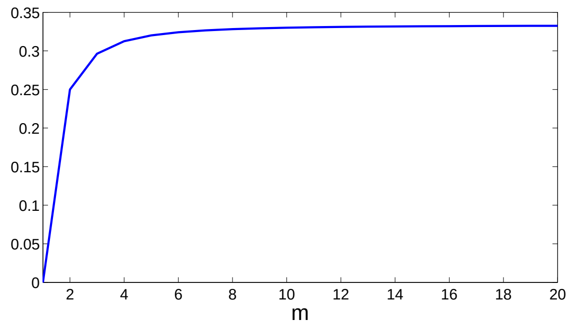

Figure 3. $\Pi(\nabla\widetilde{z})$ for $z_t = z_{t-1} + \varepsilon_t$

Remark 6 Stock data may be aggregated by skip sampling or averaging. The way how persistence is affected may differ dramatically. If the stock data at hand are integrated of order one ($r = 1$), then skip sampling results by Proposition 2*a)* in $\Pi(\nabla\dot{z}) = \Pi(\widetilde{y})$, while averaging yields $\Pi(\nabla\overline{z}) = \Pi(\nabla\widetilde{z})$. If the random walk from Remark 5 is averaged, then Figure 3 is reproduced; if the random walk is skip sampled, then the persistence is not affected: $\Pi(\nabla\dot{z}) = \Pi(\Delta z) = 0$, since all $\gamma(h) = 0$ on the right-hand side of (11). But skip sampling cannot generally be preferred to averaging. Just as in Remark 5 it holds: Whether $\Pi(\nabla\overline{z})$ (with limit 1/3) or $\Pi(\nabla\dot{z})$ (with limit 0) is closer to the persistence of the underlying process depends on the value $\Pi(\Delta z)$ characterizing $\{z_t\}$.

Finally, we turn to the effect of computing annual growth rates on persistence.

Proposition 3 *Let $\Delta z_t = y_t$, where $\{y_t\}$ satisfies Assumption 1 with autocovariances $\gamma(h)$. It then holds for $\Delta_m z_t$:*

$$\begin{aligned} \Pi(\Delta_m z) &= 1 + \frac{\Pi(\widetilde{y}) - 1}{m} \\ &\to 1 \quad \textit{as } m \to \infty \,, \end{aligned} \tag{13}$$

where $\Pi(\widetilde{y})$ is from Proposition 1.

Proof See Appendix.

Remark 7 In order to illustrate (13), let us assume that $\Delta z_t = y_t$ is white noise and $m = 12$. This means that monthly growth rates are free of persistence, $\Pi(\Delta z) = 0$. One obtains $\Pi(\widetilde{y}) = 0$ by Proposition 1, and hence by Proposition 3 that $\Pi(\Delta_{12} z) = 0.9167$. This is the same value produced by (6) for $a = 0.846$ and $b = 0$. In other words: Annual growth rates computed from a monthly random walk are as persistent as an AR(1) process with autoregressive parameter $a = 0.846$, although the seasonal growth rates are free of persistence. Such spurious persistence has been observed by Hassler and Demetrescu (2005), who argued that evidence in favour of unit roots in annual growth rates might be caused that way.

5 Empirical application

In order to illustrate the usefulness of Propositions 1 through 3 we use some textbook time series where the data generating process behind $\{z_t\}$ is assumed to be known or at least to be widely acceptable. We next aggregate (and difference) the data, predict what the persistence of the (differenced) aggregate should be like in view of our theory, and compare this value with direct estimates obtained from the aggregates. The direct estimation relies on

$$\widehat{\Pi} = 1 - \frac{\widehat{\gamma}(0)}{\widehat{\omega}^2},$$

where $\widehat{\omega}^2$ is a consistent estimator of $\omega^2 = \mathrm{LrV}(\Delta^r z_t) = \mathrm{LrV}(y_t)$. For a kernel or window w_B depending on a bandwidth B one has

$$\widehat{\omega}^2 = \widehat{\gamma}(0) + 2 \sum_{h=1}^{T-1} w_B(h)\, \widehat{\gamma}(h), \tag{14}$$

building on consistent estimates $\widehat{\gamma}(h)$. The long-run variance estimation is performed in EViews 7.2 where $\widehat{\omega}^2$ builds on a Bartlett kernel in (14) with automatic bandwidth selection according to Newey and West (1994). Throughout, we consider stock data, where aggregation may consist of averaging (with the same effect as cumulating) or skip sampling.

Random Walk

The return series on efficient markets should be serially uncorrelated. Consequently, Kirchgässner, Wolters and Hassler (2013, Ex. 1.3) demonstrate with monthly data from 1974.01 till 2011.12 that the log-differences of the exchange rate $\{e_t\}$ between the Swiss Franc and the U.S. dollar follow a white noise process. This means that $\{z_t\}$ with $z_t = \log(e_t)$ follows a random walk. In the case of a random walk, where $\{\Delta z_t\} = \{y_t\}$ is white noise, one obtains no persistence from differences: $\Pi(\Delta z) = 0$. Direct estimates of $\mathrm{Var}(y_t)$ and $\mathrm{LrV}(y_t)$ yield in this example with $T = 456$ observations

$$\widehat{\Pi}(\Delta z) = 1 - \widehat{\sigma}^2/\widehat{\omega}^2 = 0.0327 \approx 0\,,$$

which backs the random walk specification in levels.

Next, we aggregate the data to quarterly series ($m = 3$). First, Proposition 2*a)* yields for a random walk in case of skip sampling that $\Pi(\nabla\dot{z}) = 0$. This value very well explains the direct estimate obtained from $\{\nabla\dot{z}_\tau\}$, which is $\widehat{\Pi}(\nabla\dot{z}) = 0.01$. Second, when averaging a random walk we expect by Proposition 2 in connection with Remark 3 because of $\gamma(0) = \mathrm{LrV}(\Delta z)$:

$$\begin{aligned}\Pi(\nabla\overline{z}) &= 1 - \frac{\gamma(0)\ [m^2 + 2\ (1^2 + 2^2 + \cdots + (m-1)^2)]}{m^3\gamma(0)} \\ &= \frac{1}{3} - \frac{1}{3\,m^2}\,.\end{aligned}$$

Consequently, in our random walk example with $m = 3$: $\Pi(\nabla\overline{z}) = 0.30$. This value is reasonably close to the direct estimate obtained from $\{\nabla\overline{z}_\tau\}$: $\widehat{\Pi}(\nabla\overline{z}) = 0.26$. Third, we compute quarterly returns, $\Delta_3 z_t$. Theoretically, we expect from Proposition 3: $\Pi(\Delta_3 z) = 1 - 1/3 = 0.67$, see Remark 7. In very close correspondence, the direct estimate yields $\widehat{\Pi}(\Delta_3 z) = 0.68$.

In summary, the random walk assumption for the Swiss-U.S. exchange rate data yields: Persistence in the monthly log-exchange rate data is conserved upon skip sampling, but strongly exaggerated in the case of averaging; similarly, quarterly returns tremendously overdo the persistence present in monthly returns. See Table 1 for a numerical summary.

Table 1. Persistence in aggregates of exchange rate (random walk)

	$\Pi(\Delta z)$	$\Pi(\nabla \dot{z})$	$\Pi(\nabla \overline{z})$	$\Pi(\Delta_3 z)$
direct estimation	0.03	0.01	0.26	0.68
Prop.s 2 and 3	0.00	0.00	0.30	0.67

Notes: Estimates of Π from monthly data z and quarterly aggregates ($m = 3$: average $\overline{z}$ and skip sampled $\dot{z}$), and quarterly differences ($\Delta_3 z$); direct estimates from differences with Bartlett kernel and automatic bandwidth selection in (14) compared with theoretical values based on the random walk model and Propositions 2 and 3.

Stationary AR(1)

As an example for a stationary AR(1) series, $z_t = y_t$, Kirchgässner *et al.* (2013) use the monthly data on the popularity of the Christian Democratic Party in Germany. This popularity series consists of the share of voters who answered they would vote for this party if there were federal elections the following Sunday. The monthly data range from December 1970 to April 1982. Popularity is another example of stock data. We compute and compare quarterly aggregates ($m = 3$) by averaging and skip sampling.

For the monthly data $\{y_t\}$ Kirchgässner *et al.* (2013, Ex. 2.2) estimate

$$y_t = 8.053 + 0.834\, y_{t-1} + \widehat{\varepsilon}_t\,, \quad t = 1, ..., 136\,. \tag{15}$$

Example 1 yields for $a = 0.834$ and $b = 0$: $\Pi(y) = 0.91$. Direct estimates of $\mathrm{Var}(y_t)$ and $\mathrm{LrV}(y_t)$ yield

$$\widehat{\Pi}(y) = 1 - \widehat{\sigma}^2/\widehat{\omega}^2 = 0.90\,,$$

which strongly supports the AR(1) specification.

Now, the data are aggregated to quarterly series, $m = 3$. When averaging an AR(1) process from Example 1 we obtain by Remark 3

$$\Pi(\overline{y}) = 1 - \frac{(1-a)^2}{1-a^2}\left(1 + 4\,\frac{a}{3} + 2\,\frac{a^2}{3}\right)\,.$$

Consequently, in our data example with $a = 0.834$: $\Pi(\overline{y}) = 0.77$. This value is reasonably close to the direct estimate obtained from $\{\overline{y}_\tau\}$: $\widehat{\Pi}(\overline{y}) = 0.73$. Similarly, Proposition 1*a)* yields for an AR(1) process with spectral density

$$2\pi\, f(\lambda) = \frac{\sigma^2}{1 - 2a\, \cos(\lambda) + a^2}$$

that

$$\Pi(\dot{y}) = 1 - \frac{3\,(1-a^2)^{-1}}{(1-a)^{-2} + 2\,(1+a+a^2)^{-1}}\,.$$

Consequently, we predict for the popularity series with $a = 0.834$: $\Pi(\dot{y}) = 0.73$. This value well explains the direct estimate obtained from $\{\dot{y}_\tau\}$, which is $\widehat{\Pi}(\dot{y}) = 0.69$.

To summarize this example: We observe that the persistence in the monthly stock data is reduced from 0.90 to 0.73 and 0.69 in case of averaging and skip sampling to quarterly data, respectively. Taking for granted that the popularity series is generated from an AR(1) process, those reductions in persistence are well explained by Proposition 1. For a numerical summary see Table 2.

Table 2. Persistence in aggregates of party popularity, AR(1)

	$\Pi(y)$	$\Pi(\overline{y})$	$\Pi(\dot{y})$
direct estimation	0.90	0.73	0.69
Prop. 1	0.91	0.77	0.73

Notes: Estimates of Π from monthly data y and quarterly aggregates ($m = 3$: average $\overline{y}$ and skip sampled $\dot{y}$); direct estimates with Bartlett kernel and automatic bandwidth selection in (14) compared with theoretical values based on (15) and Proposition 1.

6 Summary

Temporal aggregation of time series affects the serial correlation as well as the variance. Therefore, we advocate to measure persistence in relative terms, namely as percentage of the long-run variance not due to the variance ("long-run variance difference ratio"). It turns out that persistence is inherently linked to the sampling frequency and hence inevitably affected by aggregation. Aggregation covers cumulation of flow variables, as well as averages and end-of-period aggregates in the case of stock data. On top, we quantify the effect on persistence when computing annual growth rates (seasonal differences) instead of seasonal growth rates (usual differences).

We exactly quantify the effect of aggregation and differencing in terms of the autocovariances of the underlying basic process, and approximations are available for growing aggregation level. With numerical examples, the following results are illustrated. Differencing and aggregation are not exchangeable (see Remark 4). If one cumulates a stationary flow variable, the persistence will tend to be reduced and converge to zero with growing aggregation level, see Remark 2. This is even more true when skip sampling stationary stock data (Remark 1). On the contrary, if one differences a flow variable after cumulation, the persistence will tend to 1/3 and will hence be over-estimated (under-estimated) if it is smaller (larger) in the underlying series without aggregation (Remark 5). If stock variables are averaged this has the same effect as cumulating flows (Remark 3). Skip sampling, on the other hand, has under nonstationarity the same effect as cumulation under stationarity (Remark 6). Generally, one observes that no aggregation method is always superior to the other ones, because the aggregation effect of a certain method inevitably depends on the autocorrelation structure of the data at hand. Lastly, annual growth rates as seasonal differences will exaggerate the persistence of seasonal growth rates sizeably (Remark 7).

Although it is impossible to provide general guidelines how to best aggregate in practice, we stress that our results are relevant for empirical research. With real data at hand the effect of temporal aggregation on persistence measured as long-run variance difference ratio can be explained and predicted for an arbitrary aggregation level. The explanation is exact in theory, but will be confined in practice by estimation errors of course, as we demonstrate with example series.

Appendix

A lemma

Most of the above results build on the following lemma.

Lemma 1 *Let $\{z_t\}$ solve for some natural number $r \in \{0, 1, 2, \ldots\}$ the difference equation (7), where $\{y_t\}$ satisfies Assumption 1 with autocovariances $\gamma(h)$. It then holds*

a) for finite $m \geq 2$ with $n = m - 1$:

$$\Pi(\nabla^r \widetilde{z}) = 1 - \frac{Var(y_t) \sum_{i=0}^{rn+n} b_{r,i}^2 + 2 \sum_{h=1}^{rn+n} \left[\gamma(h) \sum_{i=0}^{rn+n-h} b_{r,i} b_{r,i+h}\right]}{m^{2r+1} \, LrV(y_t)}$$

where $b_{r,0}$ through $b_{r,rn+n}$ are defined through the expansion

$$\sum_{i=0}^{rn+n} b_{r,i} \, L^i = \sum_{k=0}^{r+1} \binom{r+1}{k} \left[L \, S_{m-1}(L)\right]^k$$

with $S_m(L) = 1 + L + \cdots + L^{m-1}$;

b) and as $m \to \infty$:

$$\Pi(\nabla^r \widetilde{z}) \;\to\; 1 - \frac{1}{(2r+1)!} \sum_{k=0}^{r} (-1)^k \binom{2r+2}{k} (r+1-k)^{2r+1} \, .$$

Proof From Hassler (2011, Lemma 2) it follows $\mathrm{LrV}(\nabla^r \widetilde{z}_\tau) = m^{2r+1} \mathrm{LrV}(y_t)$, such that

$$\Pi(\nabla^r \widetilde{z}) = 1 - \frac{\mathrm{Var}(\nabla^r \widetilde{z}_\tau)}{m^{2r+1} \mathrm{LrV}(y_t)} \, , \tag{16}$$

where $\nabla^r \widetilde{z}_\tau = \left[S_m(L)\right]^{r+1} y_{m\tau}$. With $S_m(L) = 1 + L \, S_{m-1}(L)$ one obtains the given expansion $\left[S_m(L)\right]^{r+1} = \sum_{i=0}^{(r+1)(m-1)} b_{r,i} L^i$, such that result a) becomes obvious. From Hassler and Tsai (2013, Corollary 3) we know

$$\frac{Var\,(\nabla^r \widetilde{z}_\tau)}{m^{2r+1}} \to \mathrm{LrV}(y_t) \, \frac{1}{(2r+1)!} \sum_{k=0}^{r} (-1)^k \binom{2r+2}{k} (r+1-k)^{2r+1}$$

as $m \to \infty$. With (16) the proof of the lemma is complete. ■

Proof of propositions

Proposition 1 The first statement *a)* for skip sampling under $r = 0$ is covered by the multivariate results in Hassler (2013, Lemma A, Prop. 3). The result *b)* is elementary to establish in view of Lemma 1. Hence, the proof is complete.

Proposition 2 For *a)* we note the following relation between skip sampling and cumulation under differencing: skip sampling first and then differencing amounts to first differencing and then cumulating, or formally $\nabla \dot{z}_\tau = \widetilde{\Delta z}_\tau$. In fact, this result extends to $r \geq 1$:

$$\nabla^r \dot{z}_\tau = \nabla^{r-1} \widetilde{\Delta z}_\tau \,, \quad r = 1, 2, \dots \,, \tag{17}$$

i.e. differencing a skip-sampled $I(r)$ process r times coincides with differencing a cumulated $I(r-1)$ process $r-1$ times: $\Pi(\nabla^r \dot{z}) = \Pi(\nabla^{r-1} \widetilde{\Delta z})$, which yields *a)*. The result *b)* is elementary to establish in view of Lemma 1. Hence, the proof is complete.

Proposition 3 By definition it holds $\mathrm{Var}(\Delta_m z_t) = \mathrm{Var}(\widetilde{y}_\tau)$, because $\Delta_m z_t = S_m(L) y_t$ under the assumption of Proposition 3. Similarly, it holds for the spectrum $f_{\Delta_m z}(\lambda) = |S(e^{-i\lambda})|^2 f_y(\lambda)$ with $i^2 = -1$, such that $f_{\Delta_m z}(0) = m^2 f_y(0)$ or $\mathrm{LrV}(\Delta_m z_t) = m^2 \mathrm{LrV}(y_t)$. Hence, by Hassler (2011, Lemma 2), $\mathrm{LrV}(\Delta_m z_t) = m \, \mathrm{LrV}(\widetilde{y}_\tau)$, and it is straightforward to complete the proof.

References

Andrews, D. W. K. and Chen, H.-Y. (1994). Approximately median-unbiased estimation of autoregressive models. *Journal of Business & Economic Statistics*, **12**, 187–204.

Brewer, K. R. W. (1973). Some consequences of temporal aggregation and systematic sampling for ARMA and ARMAX models. *Journal of Econometrics*, **1**, 133–154.

Campbell, J. Y. and Mankiw, N. G. (1987). Are output fluctuations transitory? *Quarterly Journal of Economics*, **102**, 857–880.

Cochrane, J. H. (1988). How big is the random walk in GNP? *Journal of Political Economy*, **96**, 893–920.

Cogley, T. and Sargent, T. S. (2005). Drifts and volatilities: monetary policies and outcomes in the post WWII US. *Review of Economic Dynamics*, **8**, 262–302.

Hamilton, J. D. (1994). *Time Series Analysis*. Princeton University Press.

Hassler, U. (2011). Estimation of fractional integration under temporal aggregation. *Journal of Econometrics*, **162**, 240–247.

— (2013). Effect of temporal aggregation on multiple time series in the frequency domain. *Journal of Time Series Analysis*, **forthcoming**.

— and Demetrescu, M. (2005). Spurious persistence and unit roots due to seasonal differencing: The case of inflation rates. *Journal of Economics and Statistics*, **225**, 413–426.

— and Tsai, H. (2013). Asymptotic behavior of temporal aggregates in the frequency domain. *Journal of Time Series Econometrics*, **5**, 47–60.

Kirchgässner, G., Wolters, J. and Hassler, U. (2013). *Introduction to Modern Time Series Analysis*. Springer, 2nd edn.

Newey, W. K. and West, K. D. (1994). Automatic lag selection in covariance matrix estimation. *Review of Economic Studies*, **61**, 631–653.

Paya, I., Duarte, A. and Holden, K. (2007). On the relationship between inflation persistence and temporal aggregation. *Journal of Money, Credit and Banking*, **39**, 1521–1531.

Rossana, R. J. and Seater, J. J. (1995). Temporal aggregation and economic time series. *Journal of Business & Economic Statistics*, **13**, 441–451.

Souza, L. R. (2005). A note on Chambers's "long memory and aggregation in macroeconomic time series". *International Economic Review*, **46**, 1059–1062.

Stram, D. O. and Wei, W. W. S. (1986). Temporal aggregation in the ARIMA process. *Journal of Time Series Analysis*, **7**, 279–292.

Tiao, G. C. (1972). Asymptotic behaviour of temporal aggregates of time series. *Biometrika*, **59**, 525–531.

Tsai, H. and Chan, K. S. (2005). Temporal aggregation of stationary and nonstationary discrete-time processes. *Journal of Time Series Analysis*, **26**, 613–624.

Wei, W. W. S. (1981). Effect of systematic sampling on ARIMA models. *Communications in Statistics - Theory & Methods*, **23** (A10), 2389–2398.

— (1990). *Time Series Analysis: Univariate and Multivariate Methods.* Addison-Wesley.

Weiss, A. A. (1984). Systematic sampling and temporal aggragation in time series models. *Journal of Econometrics*, **26**, 271–282.

Working, H. (1960). A note on the correlation of first differences of averages in a random chain. *Econometrica*, **26**, 916–918.