Temporal Aggregation in the Frequency Domain: with application to fractional integration

Uwe Hassler*
Goethe University Frankfurt†

March 17, 2010

Abstract

A result characterizing the effect of temporal aggregation in the frequency domain (aliasing) is given for arbitrary stationary processes. Temporal aggregation includes here cumulation of flow variables as well as systematic (or skip) sampling of stock variables. Next, the aggregation result is applied to fractionally integrated processes. In particular, it is investigated whether typical assumptions made for semiparametric estimation and inference are closed with respect to aggregation. It turns out that they are closed with respect to cumulating time series, but not with respect to skip sampling. Finally, we discuss proposals repairing the shortcoming in case of lack of closedness.

Keywords: long memory, cumulating time series, skip sampling,

*An earlier version of this paper was written while visiting the University of California San Diego and was presented at Texas A&M University, the 3rd ETSERN Meeting, Nottingham, and the Conference on Recent Developments in Macroeconomics, Mannheim. I am grateful to Patrik Guggenberger, Joon Park, Philippe Soulier, Jim Stock and Yixiao Sun for a lot of support and very helpful comments.

†RuW, Grueneburgplatz 1, 60323 Frankfurt, Germany. Email: hassler@wiwi.uni-frankfurt.de

closedness of assumptions

JEL classification: C14 (Semiparametric and Nonparametric Methods), C22 (Time-Series Models), C82 (Methodology for Collecting, Estimating, and Organizing Macroeconomic Data)

1 Introduction

Determining inflation persistence is a prominent issue when it comes to forecasting (Stock and Watson, 2007), or when monetary policy recommendations are at stake, see e.g. Mishkin (2007). Kumar and Okimoto (2007) addressed the possibility of breaks in inflation persistence within a framework of fractional integration, which can be traced back to Hassler and Wolters (1995) or Baillie, Chung and Tieslau (1996). The effect of temporal aggregation on inflation dynamics has recently been studied by Paya, Duarte and Holden (2007). The question how aggregation and persistence interact is of interest beyond inflation, and has troubled applied economists for a long time, see Christiano, Eichenbaum and Marshall (1991) for empirical evidence in the context of the permanent income hypothesis and Rossana and Seater (1995) for a representative set of economic time series. Using fractionally integrated models, Chambers (1998) found with macroeconomic series that the empirical degree of integration may depend on the level of temporal aggregation, see also Diebold and Rudebusch (1989). In empirical finance, too, one of the core issues with respect to realized volatility is optimal sampling, see e.g. Ait-Sahalia, Mykland and Zhang (2005) or Andersen and Bollerslev (1998).

In this paper we understand by temporal aggregation both: systematic sampling (or skip sampling) of stock variables where only every pth data point is observed, and summation of flow variables where neighbouring observations are cumulated to determine the total flow. Econometricians have devoted their attention to both types of temporal aggregation for decades. Early results for autoregressive moving-average (ARMA) models were obtained by Brewer (1973) and Weiss (1984), and by Geweke (1978) for stationary dynamic regression models. A treatment of integrated (of order one) ARIMA models was provided by Wei (1981) and Stram and Wei (1986), for skip sampling and cumulating, respectively. In particular, skip sampling can be embedded in the more general problem of missing observations, see Palm and Nijman (1984) for an investigation of dynamic regression models. In the frequency domain, temporal aggregation will be accompanied by the so-called aliasing effect, which is well known under discrete-time sampling from a continuous-time process, see e.g. Sims (1971) and Hansen and Sargent (1983). In particular, the aspect of temporal aggregation and forecasting has been addressed by Lütkepohl (1987). Moreover, the potential interaction of seasonal integration and unit roots at frequency zero due to temporal aggregation was studied by Granger and Siklos (1995), see also Pons (2006).[1]

We add two aspects to this literature: a general characterization of time aggregation in the frequency domain, and an investigation how assumptions for semiparametric inference of a fractionally integrated model are affected under temporal aggregation. In greater detail our contributions are the following. First, we study the effect of temporal aggregation (cumulating flow variables or systematic skip sampling stock variables) of an arbitrary stationary process in the frequency domain. Several results that are implied or at least suggested in the literature are here explicitly collected under general conditions (Proposition 1). Second, the aggregation result is applied to fractionally integrated processes. In particular, we investigate whether typical assumptions on fractionally integrated processes, which are made in

[1]In fact, there is a literature on "span versus frequency" when it comes to testing the null hypothesis of a unit root. With the span of data held fixed, Shiller and Perron (1985) varied the number of observations by changing the frequency of data. Experimentally, it turned out that the power of unit root tests depends rather on the span of time than on the number of observations. This was reinforced theoretically with asymptotic arguments by Perron (1991): unit root tests are consistent only if the span of data is increasing. These results were established for point-sampling from a continuous-time process, which corresponds to the case of stock variables. Still, similar findings were discovered experimentally for flow variables by Choi (1992) and Ng (1995). Finally, Chambers (2004) extended Perron's (1991) asymptotic results to the case of flow variables as long as an intercept is included in the unit root regression.

the literature to obtain consistency or limiting normality of semiparametric estimators, are closed with respect to aggregation. In other words: if $\{y_t\}$ satisfies a set of assumptions $\mathcal{A}$ (which are sufficient to prove properties of some estimator or test), does the temporal aggregate fulfill $\mathcal{A}$, too? If not, then we should be worried, because in most cases there is no "true" or "natural" frequency of the data generating process (DGP), i.e. our observed data must be considered as aggregates. If they do not satisfy $\mathcal{A}$ upon aggregation, then we lose grounds for reliable inference. Third, it turns out that typical spectral assumptions made in the semiparametric long memory literature are closed with respect to cumulating and averaging the data (Proposition 2 and Corollary 1). Fourth, it is established that certain spectral assumptions are not closed with respect to skip sampling fractional integration (Proposition 3). Fifth, we discuss repair proposals to this shortcoming.

The rest of this paper is organized as follows. Section 2 treats aggregation in terms of spectral densities. In Section 3, the aggregation result is applied to fractional integration, and the effects of temporal aggregation on integration are studied in some detail. The last section contains a non-technical summary. Proofs are relegated to the Appendix.

2 Aggregation in the frequency domain

For sequences $\{a_j\}$ and $\{b_j\}$, let $a_j \sim b_j$ denote $a_j/b_j \to 1$ as $j \to \infty$, while for functions, $a(x) \sim b(x)$ is short for $a(x)/b(x) \to 1$ as $x \to 0$. Further, $a(x) = O(x^c)$ means that $a(x)\, x^{-c}$ is bounded as $x \to 0$, while $a(x) = o(x^c)$ signifies $a(x)\, x^{-c} \to 0$. Finally, let $\mathbb{Z}$ stand for the set of all integers.

2.1 Notation and assumptions

Let $\{y_t\}$, $t = 1, 2, \ldots, T$, denote some time series to be aggregated over p periods. For simplicity we assume $T = p\,N$ for some integer N, and the aggregate is constructed for the new time scale $\tau = 1, \ldots, N$. In case of stock variables, aggregation or systematic sampling means *skip sampling* where only every p'th data point is observed,

$$\dot{y}_\tau := y_{p\tau}, \quad \tau = 1, 2, \dots , \tag{1}$$

where for the rest of the paper $p \geq 2$ is a finite integer. Flow variables are aggregated by *cumulating* p neighbouring observations that do not overlap to determine the total flow over p sub-periods,

$$\begin{aligned} \widetilde{y}_\tau &:= y_{p\tau} + y_{p\tau-1} + \dots + y_{p(\tau-1)+1} \\ &= S_p(L)\, y_{p\tau}\,, \quad \tau = 1, 2, \dots , \end{aligned} \tag{2}$$

where $S_p(L) := 1 + L + \cdots + L^{p-1}$ is the moving average filter of order p. The link between the two aggregates is given by a moving average of order p where observations overlap,

$$y_t^{ma} \;:=\; S_p(L)\, y_t\,, \tag{3}$$

because $\widetilde{y}_\tau$ is obtained from skip sampling the moving average, which amounts to

$$\dot{y}_\tau^{ma} = S_p(L)\, y_{p\tau} = \widetilde{y}_\tau\,, \quad \tau = 1, 2, \dots .$$

The main result on the effect of temporal aggregation in the frequency domain holds for any stationary process $\{y_t\}$ with autocovariances $\gamma(h) = E(y_t y_{t+h}) < \infty$ and spectral density f, where for simplicity $E(y_t) = 0$. The link between the time domain (autocovariances) and the spectral density $f(\lambda)$ in the frequency domain is given by Fourier transformation for $|\lambda| \leq \pi$:

$$\begin{aligned} f(\lambda) &= \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \gamma(h) \exp(-i\,\lambda\, h)\,, \quad i^2 = -1\,, \\ \gamma(h) &= \int_{-\pi}^{\pi} f(\lambda) \exp(i\,\lambda\, h) d\lambda\,. \end{aligned}$$

Since f is an even and 2π-periodic function, the definition of the spectral density can be extended to the whole real range, and we can focus on the interval $[0, \pi]$ in the following assumption.

Assumption 1 *The process $\{y_t\}$, $t \in \mathbb{Z}$, is covariance stationary with autocovariances $\gamma(h)$ and a spectral density $f(\lambda)$ on Π, where $\Pi = [0,\pi]$ if f is well defined on the whole interval, or $\Pi = [0,\pi] \setminus \{\lambda_0\}$ if f has a pole at some frequency $\lambda_0 \in [0,\pi]$.*

We only require that f is integrable over $[0,\pi]$, although it does not have to exist everywhere. In particular, a pole at $\lambda_0 = 0$ might come from fractional integration with long memory, see (8) below. Similarly, we might allow for k poles (having e.g. so-called k-factor Gegenbauer processes in mind, see Woodward, Cheng and Gray, 1998).

2.2 Results and discussion

Let the spectral densities of the aggregates $\{\dot{y}_\tau\}$ from (1) and $\{\widetilde{y}_\tau\}$ from (2) be denoted as $\dot{f}(\lambda)$ and $\widetilde{f}(\lambda)$, respectively. The following properties are implied by related results from the literature, but seem to be explicitly collected here for the first time under conditions as general as Assumption 1.

Proposition 1 *Let $\{y_t\}$ be from Assumption 1 and assume that f is bounded at $(\lambda + 2\pi j)/p$, $j = 1, \ldots, (p-1)$. It then holds for the aggregates of $\{y_t\}$:*

a) *in case of skip sampling*

$$\dot{f}(\lambda) \;=\; \frac{1}{p}\sum_{j=0}^{p-1} f\left(\frac{\lambda + 2\,\pi\, j}{p}\right)\;;$$

b) *in case of cumulating*

$$\widetilde{f}(\lambda) = \frac{1}{p}\sum_{j=0}^{p-1} f\left(\frac{\lambda + 2\,\pi\, j}{p}\right)\phi_j(\lambda)\,,$$

where

$$\phi_j(\lambda) = \frac{\sin^2\left(\frac{\lambda}{2} + \pi\, j\right)}{\sin^2\left(\frac{\frac{\lambda}{2}+\pi\, j}{p}\right)}\,, \quad \lambda > 0\,,$$

$$\phi_0(\lambda) \sim p^2 , \quad \phi_j(\lambda) \;\sim\; \frac{\lambda^2}{4\,\sin^2\left(\frac{\pi j}{p}\right)} , \quad j = 1, \ldots, p-1 ,$$

as $\lambda \to 0$*. Further,* ϕ_j *is continuously differentiable with (*$\lambda \to 0$*)*

$$\phi_j'(\lambda) = O(\lambda) , \quad j = 0, \ldots, p-1 .$$

Proof See Appendix.

Remark A The summation over the frequencies $\frac{\lambda+2\pi j}{p}$, $j = 0, 1, \ldots, p-1$, in Proposition 1a) corresponds to the well known aliasing effect that occurs when observing a continuous-time process at discrete points in time, see e.g. Hansen and Sargent (1983), or the discussions in Bloomfield (2000, p.196) and Priestley (1981, p.224, p.506): Cycles of frequency λ/p in the original data become cycles of frequency λ upon skip sampling. The cumulated aggregate is subject to aliasing, too (Proposition 1b)), simply because $\{\widetilde{y}_\tau\}$ is constructed from a moving average through skip sampling. In this case, however, aliasing is superimposed by the factors ϕ_j due to the moving average filter.

Sometimes stock variables are aggregated by averaging over p non-overlapping observations, $\{\overline{y}_\tau\}$, such that p sub-periods are replaced by the mean of p values. Obviously this is directly connected to cumulation from (2):

$$\overline{y}_\tau := \frac{\widetilde{y}_\tau}{p}, \quad \tau = 1, 2, \ldots . \tag{4}$$

Let the spectral density of the aggregate $\{\overline{y}_\tau\}$ from (4) be denoted as $\overline{f}(\lambda)$. The results from Proposition 1b) carry over to $\overline{f}$ in the obvious way because $\overline{f}(\lambda) = \frac{\widetilde{f}(\lambda)}{p^2}$.

Next, we want again to allow for nonstationarity where the high frequency variable $\{z_t\}$ is given by integration,

$$z_t = z_0 + \sum_{i=1}^{t} y_i , \quad t = 1, 2, \ldots, T , \tag{5}$$

with $\{y_t\}$ satisfying the previous assumptions. If $\{y_t\}$ is a stationary fractionally integrated process of order d as defined in the next section, then the cumulation $\{z_t\}$ is sometimes called fractionally integrated (of order $\delta = 1 + d$) of "type I", see Marinucci and Robinson (1999) and Robinson (2005). Following Velasco (1999a), one may define a pseudo spectral density of the integrated process as $f_z(\lambda) = |1 - e^{i\lambda}|^{-2} f_y(\lambda)$, $\lambda > 0$, where $f_y(\lambda)$ belongs to $\{y_t\}$. This is just another way of stating that the spectral density of Δz_t (with $\Delta = 1 - L$, and L is the usual lag operator) is f_y: $f_{\Delta z}(\lambda) = f_y(\lambda)$, $\lambda \in \Pi$. Differencing and temporal aggregation are not exchangeable. Consider $\{\dot{z}_\tau\}$ and $\{\widetilde{z}_\tau\}$ constructed as in (1) and (2), respectively, and define the lag operator $\mathcal{L}$ that operates on the aggregate time scale τ, such that $\mathcal{L} = L^p$ with L operating on t (see e.g. Wei, 1990, Ch.16). Results for the differenced aggregates are readily available. For the skip sampled aggregate ($\tau = 1, 2, \ldots$)

$$(1 - \mathcal{L})\dot{z}_\tau = S_p(L)\, \dot{y}_\tau = \dot{y}_\tau^{ma} = \widetilde{y}_\tau\,, \tag{6}$$

and similarly for flow variables:

$$(1 - \mathcal{L})\widetilde{z}_\tau = S_p(L)\, \widetilde{y}_\tau = S_p(L)\, y_{p\tau}^{ma} = \widetilde{y_\tau^{ma}}\,.$$

Hence, differencing the aggregates results in moving averaging differences, which holds true because $(1 - \mathcal{L}) = (1 - L^p) = S_p(L)\,(1 - L)$.

3 Aggregation of fractional integration

In this section we apply the results from Proposition 1 to fractionally integrated models characterized in the frequency domain. The case of discrete-time sampling from a continuous-time long memory process has been covered by Chambers (1996). The effect of temporal cumulation with p getting large was treated in Man and Tiao (2006). Here we consider for finite p the aggregation of discrete-time processes.

3.1 Assumptions

Chambers (1998) and Souza (2005) study the fractionally integrated process $\{y_t\}$ constructed from the filter $(1-L)^{-d}$ with the usual expansion,

$$y_t \quad = \quad (1-L)^{-d}\, e_t \,, \qquad \text{with } \; -1 < d < 0.5 \,, \tag{7}$$

where $\{e_t\}$ is a stationary and invertible $I(0)$ process in the following sense.

Assumption 2 *Let $\{e_t\}$, $t \in \mathbb{Z}$, be a linearly regular, absolutely summable process,*

$$e_t = \sum_{j=0}^{\infty} \rho_j \, \varepsilon_{t-j} \,, \quad \sum_{j=0}^{\infty} |\rho_j| < \infty \,,$$

where

$$\rho(0) := \sum_{j=0}^{\infty} \rho_j \neq 0 \,, \quad \rho_0 = 1 \,,$$

and $\{\varepsilon_t\}$ is white noise with mean 0 and variance σ^2.

For the fractionally integrated process $\{y_t\}$ from (7) we know

$$f(\lambda) = |1 - e^{i\lambda}|^{-2d} f_e(\lambda) = 4^{-d} \left(\sin \frac{\lambda}{2} \right)^{-2d} f_e(\lambda)$$

with f_e denoting the spectral density of $\{e_t\}$ from Assumption 2. Note that Assumption 2 implies $0 < f_e(0) < \infty$. Equivalently (because $|1 - e^{i\lambda}|^{-2d} = \lambda^{-2d}(1 + o(1))$ fractional integration is characterized through the assumption

$$f(\lambda) = \lambda^{-2d} f_e(\lambda) \,, \quad -1 < d < 0.5 \,. \tag{8}$$

Papers on semiparametric inference of long memory typically assume that the observed process has a spectral density like in (8) where the short memory component f_e is characterized by assumptions $\mathcal{A}$ as weak as possible. We consider typical spectral assumptions next.

Assumption 3 *Let $\mathcal{A}$ be a set of assumptions for $f(\lambda) = \lambda^{-2d} f_e(\lambda)$, $d < 0.5$, including*

(A0) *f_e is bounded and bounded away from zero at frequency zero;*

(A1) *f_e has a finite first derivative f_e' in a neighboorhood $(0, \epsilon)$ of zero, and*

$$f_e'(\lambda) = O(\lambda^{-1}) \,, \quad \lambda \to 0 \,;$$

(A2) *f_e has a finite first derivative f_e' at zero.*

The first assumption **(A0)** that $f_e(0)$ is bounded and positive is minimal and common to all papers in order to identify d from (8). Next, some papers work under the assumption that f_e' exists in a neighbourhood of zero but may diverge at appropriate rate as getting close to zero, see Assumption **(A1)**. Although put slightly differently such an assumption is found in Robinson (1995a, Ass. A2) and Shimotsu and Phillips (2005, Ass. 2) when establishing consistency of the local Whittle (LW) estimator and the so-called exact LW estimator, respectively[2]. A related but slightly weaker condition is employed in Robinson (1994, Ass. 4) and Lobato and Robinson (1996, (C2)) to determine optimal spectral bandwidth rates and limiting properties of the averaged periodogram estimator, respectively. Other papers assume a stronger degree of smoothness of f_e at frequency zero in that they demand the first derivative $f_e'(0)$ to be zero or at least to be finite[3], which is our assumption **(A2)**. Hurvich, Deo and Brodsky (1998) for instance assume $f_e'(0) = 0$ when deriving the asymptotic mean squared error and limiting distribution of the log-periodogram regression (LPR) by Geweke and Porter-Hudak (1983), while Andrews and Guggenberger (2003) discuss properties of a bias-reduced version under a smoothness assumption requiring $f_e'(0)$ to exist, see also Guggenberger and Sun (2006). Under similar assumptions Andrews and Sun (2004) improved on the LW estimator.

We wish to investigate which set of assumptions is closed in the following sense.

[2]See also the assumption $|f_e'(\lambda)| \leq c\,\lambda^{-1}$ for $\lambda > 0$ in Moulines and Soulier (1999, Ass. 2), and similar although slightly weaker in Soulier (2001, Ass. 1).

[3]Note that e_t from Assumption 2 satisfies **(A2)**, since $e_t = \sum_{j=0}^{\infty} \rho_j \varepsilon_{t-j}$ with absolutely summable MA coefficients implies $f_e'(0) = 0$.

Definition 1 *A set of assumptions on some process* $\{y_t\}$ *is called closed with respect to temporal aggregation (skip sampling, averaging or cumulating), if* $\{\dot{y}_\tau\}$, $\{\overline{y}_\tau\}$ *or* $\{\widetilde{y}_\tau\}$, *respectively, satisfy the same set of assumptions for any finite positive integer* $p \geq 2$, *too.*

For practical purposes procedures with properties established under assumptions that are closed with respect to aggregation are desirable, because in most practical situations a "true" frequency of the DGP is not known or does not exist. Most economic and financial time series have to be considered as aggregates. And a statistical procedure relying on assumptions $\mathcal{A}$ cannot be safely applied to an aggregate, unless $\mathcal{A}$ is closed with respect to temporal aggregation.

Since the following results are obtained under temporal aggregation we need spectral assumptions for $\lambda > 0$ due to the aliasing effect. We require that the spectral density is "well behaved" at multiples of the so-called Nyquist frequency $2\pi/p$, see Proposition 1. The usual long memory literature not addressing the aggregation issue does not need Assumption 4. This stronger assumption reads as follows.

Assumption 4 *The process* $\{y_t\}$ *from Assumption 1 has a spectral density* $f(\lambda)$, *which at frequencies* $2\pi\, j/p$, $j = 1, \ldots, (p-1)$, *is bounded, bounded away from zero and continuously differentiable with derivative* f'.

3.2 Cumulating flows and averaging stocks

In case of flow variables with (8), we have the following closedness properties.

Proposition 2 *Let* $\{y_t\}$ *be a stationary and invertible* $(-1 < d < 0.5)$ *process with spectral density as in (8) with Assumption 4. It then holds with respect to cumulated sampling of flow variables:*

a) ***(A0)*** *is closed for all* $d > -1$;

b) ***(A1)*** *is closed for all* $d > -1$;

c) ***(A2)*** *is closed if and only if* $d \geq -0.5$.

Proof See Appendix.

With $\overline{f}(\lambda) = \widetilde{f}(\lambda)/p^2$ it is obvious that the above assumptions are closed with respect to averaging stock variables according to (4) under the conditions of Proposition 2. We have the following corollary.

Corollary 1 *Let* $\{y_t\}$ *satisfy the assumptions of Proposition 2. It then holds with respect to non-overlapping averaging of stock variables,* $\{\overline{y}_\tau\}$*:*

a) ***(A0)*** *is closed for all* $d > -1$*;*

b) ***(A1)*** *is closed for all* $d > -1$*;*

c) ***(A2)*** *is closed if and only if* $d \geq -0.5$.

Proof Obvious.

The proof of Proposition 2 relies on a decomposition of the spectral density into a part due to integration and a short memory component $\widetilde{\varphi}$,

$$\widetilde{f}(\lambda) = \lambda^{-2d}\widetilde{\varphi}(\lambda)\,.$$

The latter turns out to be so smooth at the origin that **(A0)** and **(A1)** or even **(A2)** hold for $\widetilde{\varphi}$ as long as they hold for f_e from (8). Hence these conditions are closed in the sense of Definition 1 with respect to aggregation of flow variables and averaging of stock variables as long as $d \geq -0.5$. The following remark stresses implications for applied work.

Remark D In particular, Proposition 2a) confirms the finding by Chambers (1998), Hwang (2000), and Souza (2005) that the order of fractional integration is maintained under cumulated aggregation of flow variables. Similarly, assumptions about the spectral slope behaviour are inherited by the aggregate from the original series, as required for semiparametric estimators like log-periodogram regression or local Whittle estimation. In fact, Souza (2007) proved under mildly stronger assumptions than our Assumption 4 that both estimators retain their limiting normal distribution when cumulating flow data as long as the bandwidth parameter is adjusted accordingly. More strongly, Souza (2008) found that even for fixed bandwidths the limiting distributions remain unchanged upon cumulating flow data. Obviously, the results by Souza (2007, 2008), apply to non-overlapping averages of stock data (Corollary 1), too. Further, Ohanissian, Russell and Tsay (2008) proposed a test on whether there is true long memory (i.e. fractional integration) or not, that builds on comparing differences of estimators obtained from different levels of cumulation. Under the null hypothesis the difference between the estimators from the different levels of aggregation will vanish. The proof in Ohanissian *et al.* (2008) heavily relies on Soulier (2001); not surprisingly, his Assumption 1 (very similar to our **(A1)**) is closed with respect to cumulated sampling.

3.3 Systematic sampling

The results for skip sampling very much differ from the previous ones. The reason can be seen from Proposition 1a): The effect of the frequencies $\frac{\lambda+2\pi j}{p}$ on $\dot{f}$ is not negligible in a vicinity of zero. The consequences on fractional integration are given in the next proposition.

Proposition 3 *Let $\{y_t\}$ be a stationary and invertible ($-1 < d < 0.5$) process with spectral density as in (8) with Assumption 4 and* ***(A0)****. In case of skip sampling it holds that*

a) *the spectral density is given as*

$$\dot{f}(\lambda) = \lambda^{-2d}\dot{\varphi}(\lambda) \quad \text{with} \quad \dot{\varphi}(\lambda) \sim \varphi_0 + \varphi_1 \lambda^{2d} \text{ as } \lambda \to 0\,, \tag{9}$$

where $\varphi_0 = p^{2d-1} f_e(0)$, $0 < \varphi_1 < \infty$;

b) ***(A0)*** *is closed if and only if $d \geq 0$;*

c) ***(A1)*** *is closed if and only if* $d \geq 0$*;*

d) ***(A2)*** *is not closed for all values of* d*.*

Proof See Appendix.

With three remarks we discuss consequences for applied work.

Remark E With the short memory component $\dot{\varphi}$ of the spectral density in (9) being of order λ^{2d} after skip sampling, **(A0)** is not closed for negative d. For $d < 0$ with $f(0) = 0$, we get $\dot{f}(0) = \varphi_1 > 0$, and hence the aggregate $\dot{y}_\tau$ loses the spectral properties of $y_t \sim I(d)$ with $d < 0$, which corrects differing claims made in Chambers (1998) and Hwang (2000), and confirms the point made by Souza (2005): if $d < 0$, the skip sampled aggregate does not inherit the spectral properties of the original series at frequency zero. This is a remarkable result, since fractional processes are known to be self-similar in that stretching the time scale leaves distributional properties unchanged upon rescaling the process, see e.g. Mandelbrot and van Ness (1968). In fact, for ARFIMA processes it holds for $|d| < 0.5$ that

$$\dot{\gamma}(h) = E(\dot{y}_\tau \, \dot{y}_{\tau+h}) = \gamma(p\, h) \sim C\,(ph)^{2d-1}\,, \quad h \to \infty$$

for some constant C. Hence, the hyperbolic decay of the autocovariance is inherited by the skip sampled process irrespective of the sign of d, while the power law in (8) is lost for $d < 0$. How relevant is this lack of closedness in the frequency domain in practice? Assume that some variable z_t is integrated of order δ between 0.75 and 1, such that the usual limiting normality of estimators from the log-periodogram regression or local Whittle breaks down, see Velasco (1999a,b). Consequently, people have worked with differences $y_t = z_t - z_{t-1}$ in applied papers, where y_t is $I(d)$ with $d = \delta - 1 < 0$. What happens in case of skip sampling? In the unlucky case where you skip sample the differences $y_t \sim I(d)$ and try to estimate d from $\dot{y}_\tau$, any frequency-domain based estimate $\widehat{d}$ will tend towards zero even asymptotically because $\dot{y}_\tau$ behaves like $I(0)$ with $0 < \dot{f}(0) < \infty$. While this seems worrisome at first glance, there is a simple and obvious solution to the problem. One simply should aggregate the level, $\dot{z}_\tau$, before differencing. As we know from (6), the differences thereof behave like $\widetilde{y}_\tau$. Consequently, the order $d+1$ of z_t can be discovered consistently from $\dot{z}_\tau - \dot{z}_{\tau-1} = \widetilde{y}_\tau$, see Proposition 2 and Remark D.

Remark F Skip sampling preserves Assumption **(A1)** for $d \geq 0$. Therefore, the condition by Robinson (1995a, Ass. A2) used to prove consistency of the local Whittle estimator continues to hold after systematic sampling. Further, we may conclude from (9) that limiting normality of the LW estimator is guaranteed after skip sampling, since $\dot{f}(\lambda) \sim \varphi_0 \lambda^{-2d}(1 + O(\lambda^{2d}))$ as required for Robinson (1995a, Ass. A1'). However, the order of integration d will affect the required rate of divergence of the bandwidth m, say. Because of Robinson (1995a, Ass. A4') we require in the light of (9) as the sample size T goes to infinity:

$$\frac{1}{m} + \frac{m^{1+4d}(\log m)^2}{T^{4d}} \ \rightarrow \ 0\,.$$

Consequently, the bandwidth for the LW estimator has to satisfy

$$m = o\left(T^{4d/(1+4d)}\right)\,. \tag{10}$$

For values of d close to zero this means a very slow divergence of m, and hence a very slow convergence to the limiting distribution since the variance is proportional to $1/m$.

Remark G The short memory component $\dot{\varphi}$ of the aggregated spectral density in (9) displays an unbounded derivative at the origin for all $d < 0.5$ even if $f_e'(0)$ is finite, and consequently **(A2)** is never closed. This means that sufficient conditions for consistency or limiting normality of the log-periodogram regression made by Hurvich *et al.* (1998) or Andrews and Guggenberger (2003) do not hold upon systematic sampling, which sheds some doubt on the use of the LPR in applied work. Notice, however, there is a trimmed version of the LPR by Robinson (1995b), where trimming means that the first ℓ harmonic frequencies are omitted from the regression. The assumptions in Robinson (1995b, Ass. 1 and 2) correspond to Robinson (1995a, Ass. 1' and 2), and are hence closed under skip sampling, see Remark F. To ensure limiting normality of the trimmed LPR, Robinson (1995b, Ass. 6) requires in the light of (9)

$$\frac{m^{1/2}\log m}{\ell} + \frac{\ell\,(\log T)^2}{m} + \frac{m^{1+1/4d}}{T} \;\to\; 0\,,$$

which obviously implies (10). While m has again to diverge very slowly for small values of d, the trimming parameter ℓ has to diverge faster than $\sqrt{m}$, which makes appropriate choices of ℓ and m a delicate matter in practice.

3.4 Perturbed processes and LMSV

To shed further light on the effect of skip sampling it is elucidating to relate to a different strand of the literature. Let $\{x_t\}$ be a fractionally integrated process $\{y_t\}$ perturbed by some $I(0)$ process $\{u_t\}$,

$$x_t = y_t + u_t\,, \tag{11}$$

where we assume that $\{u_t\}$ is independent of the unobservable process $\{y_t\}$. Given $\{y_t\}$ is fractionally integrated with (8) it holds in the frequency domain

$$f_x(\lambda) = \lambda^{-2d} f_e(\lambda) + f_u(\lambda) = \lambda^{-2d}\varphi(\lambda)$$

where the short memory component of the observable $\{x_t\}$ becomes

$$\begin{aligned} \varphi(\lambda) &= f_e(\lambda) + f_u(\lambda)\,\lambda^{2d} \\ &\sim c_0 + c_1\,\lambda^{2d}\,, \quad \lambda \to 0\,, \end{aligned} \tag{12}$$

with $c_0 = f_e(0)$ and $c_1 = f_u(0)$. For $0 < d$, the perturbed process $\{x_t\}$ is fractionally integrated of order d where the short memory component behaves like in case of skip sampling, cf. (9): skip sampling has in the frequency domain the same effect on long memory as adding noise. Therefore, methods tailored to the estimation of d from $\{x_t\}$ in (11) are candidates for the estimation of d from skip sampled long memory series. For that reason, a short and informal review of related work is provided to close down this section.

Most papers dealing with perturbed fractional integration (also called "long memory plus noise") are related to the so-called long memory stochastic volatility model (LMSV) introduced by Breidt, Crato and de Lima (1998) or the FIEGARCH model by Bollerslev and Mikkelsen (1996). Such volatility models assume for return processes $\{r_t\}$ that

$$\log r_t^2 = \mu + y_t + \varepsilon_t \,, \tag{13}$$

where the perturbation term $\{\varepsilon_t\}$ is white noise. Sun and Phillips (2003) considered the more general model (11) under Gaussianity. They proposed an improved nonlinear version of the LPR estimator that accounts explicitly for the effect of perturbation. The bandwidth m has to obey $m = o\left(T^{8d/(8d+1)}\right)$, which is a bit less stringent than our condition (10). The same condition as (10) has been established by Arteche (2004) studying the model (11) without the assumption of Gaussianity for the LW estimator. Hurvich and Ray (2003) proposed a modification of the LW estimator adjusting explicitly for the noise effect of model (13); further refinements are provided by Hurvich, Moulines and Soulier (2005) in that correlation between y_t and ε_t is allowed for. Finally, it should be noted that the so-called broadband log-periodogram regression by Moulines and Soulier (1999) remains valid for a Gaussian LMSV model, see Iouditsky, Moulines and Soulier (1999).

4 Concluding remarks

We characterize effects of cumulating flow variables as well as systematic sampling stock variables or averaging non-overlapping stock data of an arbitrary stationary process in the frequency domain (Proposition 1). The results are applied to fractionally integrated processes. In particular, we investigate whether typical assumptions on fractionally integrated processes, which are made in the literature to justify statistical semiparametric inference, are closed with respect to aggregation. That is we study whether assumptions that hold for high-frequency data continue to hold for temporal aggregates, such that semiparametric methods like the log-periodogram regression or the local Whittle estimator are justified for aggregates, too. It turns out that typical spectral assumptions made for semiparametric estimation are closed with respect to cumulating flow variables or averaging non-overlapping stocks (Proposition 2 and Corollary 1). Hence, semiparametric procedures may be safely applied to those aggregates (Remark D).

In case of skip sampling fractionally integrated stock variables, matters are more complicated, and certain properties that hold for the high-frequency data can not be maintained for the aggregate (Proposition 3). First, if $d < 0$, the order of integration is lost upon skip sampling, and the aggregate behaves like $I(0)$ in the frequency domain. Consequently, frequency domain based methods to estimate d from the aggregate are not consistent. To repair this shortcoming it is suggested not to estimate d from aggregated differences, but from differences of aggregated data (Remark E). Second, if $d > 0$, the spectral properties are such that consistency and asymptotic normality of the log-periodogram regression are not guaranteed for skip sampled data (Remark G). Although the local Whittle estimation will remain asymptotically normal with skip sampled observations, the bandwidth has to be chosen as very small for this to hold true (Remark F). Therefore, our results suggest not to perform a long memory analysis with skip sampled stock variables. If one wants to work with aggregates of stocks, the preferred method of aggregation clearly is averaging non-overlapping observations (Remark D).

Shortcomings when estimating long memory under skip sampling may be alleviated using approaches tailored to cope with perturbed fractional integration or LMSV. An investigation how fruitful the many different procedures are in the presence of skip sampled long memory is beyond the present paper and has to be left for future research.

Appendix

Proof of Proposition 1

The proof of **a)** is similar to the proof in Bloomfield (2000, p.196) for the continuous-time case. The first result in **b)** follows because $\widetilde{y}_\tau$ is obtained from skip sampling an overlapping moving average, see e.g. Priestley (1981, p.268). Further, the $\phi_j(\lambda)$ are proportional to the so-called Fejer kernel of order p evaluated at $(\lambda+2\pi j)/p$ (see e.g. Priestley, 1981, p.400). Elementary calculus provides the behavior of $\phi_j(\lambda)$ and their derivatives $\phi_j'(\lambda)$ as $\lambda \to 0$, which completes the proof.

Proof of Proposition 2

Proposition 1b) provides under (8)

$$\widetilde{f}(\lambda) = \lambda^{-2d} p^{2d-1} f_e\left(\frac{\lambda}{p}\right) \phi_0(\lambda) + R(\lambda)$$

where

$$R(\lambda) = \frac{1}{p} \sum_{j=1}^{p-1} f\left(\frac{\lambda + 2\pi j}{p}\right) \phi_j(\lambda)\,.$$

Under Assumption 4 we obtain for $R(\lambda)$ and its derivative from Proposition 1b)

$$R(\lambda) = O(\lambda^2)\,,\ R'(\lambda) = O(\lambda)\,, \quad \lambda \to 0\,.$$

Consequently, the spectral density can be decomposed into a part due to integration and a short memory component $\widetilde{\varphi}$, $\widetilde{f}(\lambda) = \lambda^{-2d}\widetilde{\varphi}(\lambda)$. The latter is given by

$$\widetilde{\varphi}(\lambda) \quad = \quad p^{2d-1} f_e\left(\frac{\lambda}{p}\right) \phi_0(\lambda) + \lambda^{2d} R(\lambda)\,. \tag{14}$$

With the derivative R' it further holds under Assumption 4

$$\widetilde{\varphi}'(\lambda) \quad = \quad p^{2d-1} f_e'\left(\frac{\lambda}{p}\right) \frac{\phi_0(\lambda)}{p} + O(\lambda) + O(\lambda^{2d+1})\,.$$

If **(A0)** holds for f_e, than $\widetilde{\varphi}$ from (14) is finite if and only if $d \geq -1$, which proves **a)**. The results **b)** and **c)** are verified analogously by studying the behaviour of $\widetilde{\varphi}'$ given f_e' satisfies **(A1)** or **(A2)**, respectively.

Proof of Proposition 3

a) Proposition 1 yields

$$\dot{f}(\lambda) \;=\; \frac{1}{p}\left[\left[\frac{\lambda}{p}\right]^{-2d} f_e\left(\frac{\lambda}{p}\right) + \sum_{j=1}^{p-1} f\left(\frac{\lambda + 2\,\pi\,j}{p}\right)\right], \tag{15}$$

such that

$$\dot{\varphi}(\lambda) = p^{2d-1} f_e\left(\frac{\lambda}{p}\right) + \frac{\lambda^{2d}}{p}\sum_{j=1}^{p-1} f\left(\frac{\lambda + 2\,\pi\,j}{p}\right). \tag{16}$$

Hence, φ_1 becomes

$$0 < \varphi_1 = p^{-1}\sum_{j=1}^{p-1} f\left(\frac{2\,\pi\,j}{p}\right) < \infty\,,$$

and the definition of φ_0 is obvious. The statement follows from (16) under Assumption 4.

b) Under Assumption **(A0)** it follows with Assumption 4 from (9) that $\dot{\varphi}(0)$ is finite if and only if $d \geq 0$; which proves the result.

c) The derivative of $\dot{\varphi}$ is

$$\dot{\varphi}'(\lambda) = p^{2d-2} f_e'\left(\frac{\lambda}{p}\right) + 2d\,\varphi_1 \lambda^{2d-1} + O(\lambda^{2d})\,. \tag{17}$$

Given the derivative of f_e' satisfies Assumption **(A1)** one obtains:

$$\dot{\varphi}'(\lambda) = O(\lambda^{-1}) + O(\lambda^{2d-1}) + O(\lambda^{2d}) = O(\lambda^{\min(-1,2d-1)})\,;$$

which proves the result.

d) For f_e' satisfying Assumption **(A2)** we obtain directly from (17) that $\dot{\varphi}'(\lambda)$ diverges since $d < 0.5$. This completes the proof.

References

Ait-Sahalia, Y., Mykland, P. A. and Zhang, P. A. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, **18**, 351–416.

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

Andrews, D. W. K. and Guggenberger, P. (2003). A bias-reduced log-periodogram regression estimator for the long-memory parameter. *Econometrica*, **71**, 675–712.

— and Sun, Y. (2004). Adaptive local polynomial whittle estimation of long-range dependence. *Econometrica*, **72**, 569–614.

Arteche, J. (2004). Gaussian semiparametric estimation in long memory in stochastic volatility and signal plus noise models. *Journal of Econometrics*, **119**, 131–154.

Baillie, R. T., Chung, C.-F. and Tieslau, M. A. (1996). Analysing inflation by the fractionally integrated ARFIMA-GARCH model. *Journal of Applied Econometrics*, **11**, 23–40.

Bloomfield, P. (2000). *Fourier Analysis of Time Series: An Introduction.* Wiley, 2nd edn.

Bollerslev, T. and Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics*, **73**, 151–184.

Breidt, F. J., Crato, N. and de Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, **83**, 325–348.

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

Chambers, M. J. (1996). The estimation of continuous parameter long-memory time series models. *Econometric Theory*, **12**, 374–390.

— (1998). Long memory and aggregation in macroeconomic time series. *International Economic Review*, **39**, 1053–1072.

— (2004). Testing for unit roots with flow data and varying sampling frequency. *Journal of Econometrics*, **119**, 1–18.

Choi, I. (1992). Effects of data aggregation on the power of tests for a unit root. *Economics Letters*, **40**, 397–401.

Christiano, L. J., Eichenbaum, M. and Marshall, D. (1991). The permanent income hypothesis revisited. *Econometrica*, **59**, 397–423.

Diebold, F. X. and Rudebusch, G. D. (1989). Long memory and persistence in aggregate output. *Journal of Monetary Economics*, **24**, 189–209.

Geweke, J. (1978). Temporal aggregation in the multiple regression model. *Econometrica*, **46**, 643–661.

— and Porter-Hudak, S. (1983). The estimation and application of long memory time series models. *Journal of Time Series Analysis*, **4**, 221–238.

Granger, C. W. J. and Siklos, L. (1995). Systematic sampling, temporal aggregation, seasonal adjustment, and cointegration: Theory and evidence. *Journal of Econometrics*, **66**, 357–370.

Guggenberger, P. and Sun, Y. (2006). Bias-reduced log-periodogram and whittle estimation of the long-memory parameter without variance inflation. *Econometric Theory*, **22**, 863–912.

Hansen, L. P. and Sargent, T. J. (1983). The dimensionality of the aliasing problem in models with rational spectral densities. *Econometrica*, **51**, 377–387.

Hassler, U. and Wolters, J. (1995). Long memory in inflation rates: International evidence. *Journal of Business & Economic Statistics*, **13**, 37–45.

Hurvich, C. M., Deo, R. and Brodsky, J. (1998). The mean squared error of geweke and porter-hudak's estimator of the memory parameter of a long-memory time series. *Journal of Time Series Analysis*, **19**, 19–46.

—, Moulines, E. and Soulier, P. (2005). Estimating long memory in volatility. *Econometrica*, **73**, 1283–1328.

— and Ray, B. (2003). The local whittle estimator of long-memory stochastic volatility. *Journal of Financial Econometrics*, **1**, 445–470.

Hwang, S. (2000). The effects of systematic sampling and temporal aggregation on discrete time long memory processes and their finite sample properties. *Econometric Theory*, **16**, 347–372.

Iouditsky, A., Moulines, E. and Soulier, P. (1999). Adaptive estimation of the fracional differencing coefficient. *Bernoulli*, **7**, 699–731.

Kumar, M. S. and Okimoto, T. (2007). Dynamics of persistence in international inflation rates. *Journal of Money, Credit and Banking*, **39**, 1457–1479.

Lobato, I. N. and Robinson, P. M. (1996). Averaged periodogram estimation of long memory. *Journal of Econometrics*, **73**, 303–324.

Lütkepohl, H. (1987). *Forecasting Aggregated Vector ARMA Processes*. Springer.

Man, K. S. and Tiao, G. C. (2006). Aggregation effect and forecasting temporal aggregates of long memory process. *International Journal of Forecasting*, **22**, 267–281.

Mandelbrot, B. B. and van Ness, J. W. (1968). Fractional brownian motion, fractional noises and applications. *SIAM Review*, **10**, 422–437.

Marinucci, D. and Robinson, P. M. (1999). Alternative forms of fractional brownian motion. *Journal of Statistical Planning and Inference*, **80**, 111–122.

Mishkin, F. S. (2007). Inflation dynamics. *International Finance*, **10**, 317–334.

Moulines, E. and Soulier, P. (1999). Broadband log-periodogram regression of time series with long-range dependence. *The Annals of Statistics*, **27**, 1415–1439.

Ng, S. (1995). Testing for unit roots in flow data sampled at different frequencies. *Economics Letters*, **47**, 237–242.

Ohanissian, A., Russell, J. R. and Tsay, R. S. (2008). True or spurious long memory? a new test. *Journal of Business & Economic Statistics*, **26**, 161–175.

Palm, F. C. and Nijman, T. E. (1984). Missing observations in the dynamic regression model. *Econometrica*, **52**, 1415–1435.

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.

Perron, P. (1991). Test consistency with varying sampling frequency. *Econometric Theory*, **7**, 341–368.

Pons, G. (2006). Testing monthly seasonal unit roots with monthly and quarterly information. *Journal of Time Series Analysis*, **27**, 191–210.

Priestley, M. B. (1981). *Spectral Analysis and Time Series, Vol. 1.* Academic Press.

Robinson, P. M. (1994). Rates of convergence and optimal bandwidth in spectral analysis of processes with long range dependence. *Probability Theory and Related Fields*, **99**, 443–473.

— (1995a). Gaussian semiparametric estimation of long range dependence. *Annals of Statistics*, **23**, 1630–1661.

— (1995b). Log-periodogram regression of time series with long range dependence. *Annals of Statistics*, **23**, 1048–1072.

— (2005). The distance between rival nonstationary fractional processes. *Journal of Econometrics*, **128**, 283–399.

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

Shiller, R. J. and Perron, P. (1985). Testing the random walk hypothesis. *Economics Letters*, **18**, 381–386.

Shimotsu, K. and Phillips, P. C. B. (2005). Exact local whittle estimation of fractional integration. *The Annals of Statistics*, **33**, 1890–1933.

Sims, C. A. (1971). Discrete approximations to continuous time distributed lags in econometrics. *Econometrica*, **39**, 545–563.

Soulier, P. (2001). Moment bounds and central limit theorem for functions of gaussian vectors. *Statistics & Probability Letters*, **54**, 193–203.

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

— (2007). Temporal aggregation and bandwidth selection in estimating long memory. *Journal of Time Series Analysis*, **28**, 701–722.

— (2008). Why aggregate long memory time series? *Econometric Reviews*, **27**, 298–316.

Stock, J. H. and Watson, M. W. (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking*, **39**, 3–33.

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

Sun, Y. and Phillips, P. C. B. (2003). Nonlinear log-periodogram regression for perturbed fractional processes. *Journal of Econometrics*, **115**, 355–389.

Velasco, C. (1999a). Gaussian semiparametric estimation of non-stationary time series. *Journal of Time Series Analysis*, **20**, 87–127.

— (1999b). Non-stationary log-periodogram regression. *Journal of Econometrics*, **91**, 325–372.

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.

Woodward, W. A., Cheng, Q. C. and Gray, H. L. (1998). A k-factor garma long-memory model. *Journal of Time Series Analysis*, **19**, 485–504.