Temporal Aggregation of Univariate Linear Time Series Model

A. Silvestrini and D. Veredas

Discussion Paper 2005-44

Département des Sciences Économiques de l'Université catholique de Louvain

UCL

CORE DISCUSSION PAPER
2005/59

TEMPORAL AGGREGATION OF UNIVARIATE LINEAR TIME SERIES MODELS

ANDREA SILVESTRINI[1] and DAVID VEREDAS[2]

August 2005

Abstract

In this paper we feature state-of-the-art econometric methodology of temporal aggregation for univariate linear time series, namely ARIMA-GARCH models. We present a unified overview of temporal aggregation techniques for this broad class of processes and we explain in detail, although intuitively, the technical machinery behind the results. An empirical application with Belgian public deficit data illustrates the main issues.

Keywords: Temporal aggregation, ARIMA, GARCH, seasonality.
JEL classification: C10, C22, C43.

[1]Dipartimento di Economia, Finanza e Statistica, Università di Perugia, Italy and CORE, Université catholique de Louvain, Belgium.

[2]ECARES, Université Libre de Bruxelles and CORE, Université catholique de Louvain, Belgium.

Corresponding address: David Veredas, ECARES, Université Libre de Bruxelles. CP 114. 50, avenue Roosevelt, B - 1050 Bruxelles, Belgium. Tel: +32 2 650 42 18 Fax: +32 2 650 44 75. Email: dveredas@ulb.ac.be

The authors would like to thank Pierluigi Daddi, Christian Hafner, Luc Van Meensel and Bas Werker for helpful comments and suggestions. The first author gratefully acknowledges financial support from Università di Perugia. This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 INTRODUCTION

Broadly speaking, any economic variable is indexed by time. When we say *let y_t be the observation at time t*, we are implicitly assuming that the economic variable y is observed at frequency t, that is every t periods. The choice of the frequency is often given by economic arguments. But most of the times the choice of the frequency is somehow subjective. For example, when studying financial returns, should we consider daily returns, weekly or hourly?

The choice of the frequency influences the estimation results. Given the same economic or econometric model, estimation results are different for each frequency. However, it is clear that the estimated models for different frequencies should be related. A model for quarterly data should be related to a model for annual data, as the latter is a temporal aggregation of the former along the year. Therefore, not only the annual *data* is a function of the quarterly data, but also the annual *model* is a function of the quarterly model. And the quarterly estimated model is richer, information wise, because the number of observations used for estimation is four times larger than for the annual model. As a consequence, it makes sense to think that the annual model should not be estimated from annual data but rather inferred from the quarterly model.

The way in which these two models interact is the subject of this paper. In a univariate linear times series context, that is dealing with ARIMA-GARCH models, we explain how to infer the temporally aggregated model (at the low frequency) from the disaggregate one (at the high frequency). Temporal aggregation has been extensively studied in the econometric literature since the last thirty-five years and general conditions obtained in terms of order conditions (i.e. lag length), parameters estimation, asymptotic behaviour, etc. A selected literature on the effect of temporal aggregation on model structure includes Amemiya and Wu (1972), Brewer (1973), Wei (1978), Weiss (1984), Stram and Wei (1986), Drost and Nijman (1993), among others. Nonetheless, to our knowledge, a complete and up-to-date survey of the whole methodology is currently unavailable. For this reason, our aim is to present a unified overview of temporal aggregation techniques for a broad class of univariate linear time series processes. Given that ARIMA-GARCH models are the bread and butter of times series econometrics, analyzing, understanding and summarizing the consequences of temporal aggregation on model structure and parameters estimation is an issue of great theoretical and empirical relevance.

In a nutshell, deriving the low frequency model from the high frequency model involves two stages. First, ARIMA-GARCH models are specified in terms of lag polynomials whose orders have to be chosen. As we shall see, the technique of temporal aggregation permits to infer the orders of the low frequency model (i.e. annual) from those of the high frequency model (i.e. quarterly). For instance, we can answer to questions like: if the high frequency model is an ARMA(1,1), is the low frequency model an ARMA(1,1) as well? If not, which one? Second, once the orders are inferred, we recover the parameters of the low frequency model from the high frequency ones, rather than estimating them. Therefore, the parameters of the low frequency model incorporate all the information content of the high frequency data.

In general, the way in which variables aggregate may take different forms. There are indeed two aggregation schemes that are often found in economics: *stock* and *flow*.[1] *Stock* refers to aggregation in which the low frequency random variable is the result of sampling every k periods from the high frequency variable. For instance, annual observations may be obtained sampling every four periods quarterly observations. *Flow* refers to aggregation in which the low frequency variable is the sum every k periods of the high frequency variable. So that the annual observations are the sum of the quarterly observations every four periods. Some down to earth examples. Rates and indexes, such as interest rate, unemployment rate or CPI are stock variables. Direct economic measures, such as GDP, public deficit or financial returns are examples of flow variables.

[1] They are also known as systematic sampling and temporal aggregation, respectively.

The analysis of temporal aggregation starts up with the seminal article of Amemiya and Wu (1972). It shows that, if the original data are generated by an AR model of order p, the non-overlapping temporally aggregated sequence follows an AR model of order p with an MA residuals structure, regardless the aggregation frequency. Tiao (1972) and Amemiya and Wu (1972) study the issue of information loss due to incomplete sampling (i.e. sampling aggregate data instead than disaggregate data). In their articles, these authors compare predictors based on complete and incomplete samples through the variance of their prediction errors. They show that the optimal predictor built from the incomplete sample performs remarkably well with respect to the optimal predictor built from the complete sample. These findings are corroborated via numerical analysis.

Brewer, one year later, presents a generalization of the results obtained by Amemiya and Wu for pure AR models to a wider class of models. He discusses flow and stock aggregation for ARMA models with exogenous variables (called ARMAX models). Moreover, the important concept of *stability* under loss of data is introduced: a model is stable, in this sense, if additional temporal aggregation delivers a model with the same polynomial orders. In 1978, Wei derives the model structure for temporally aggregated data assuming, as underlying process, a seasonal time series in disaggregate form. He shows that, if the frequency of aggregation is the same as the seasonal frequency, the aggregate model reduces to a model without seasonality.

All the contributions so far cited deal with stationary processes. Weiss (1984), on the contrary, discusses flow and stock aggregation schemes for ARIMA models, in terms of aggregate autoregressive and moving average polynomial orders, once these schemes have been applied to the original disaggregate process. Within the large literature on temporal aggregation for ARIMA processes, Stram and Wei (1986) narrow their analysis focusing on the relationship between the autocovariance of the original disaggregate series and its aggregate counterpart. Starting from the disaggregate autocovariance function, they compute and explain the form of the aggregate autocovariance function.

More recently, Drost and Nijman (1993) derive the order conditions for temporally aggregated univariate GARCH processes. They show that, in the univariate case, only the parameters of the process followed by the flow variable depend on the disaggregate variable's fourth moment, while in the stock case there is dependence only up to the second moment. Thus, in this context, flow and stock aggregation schemes produce different outcomes. These are relevant findings since, due to their simplicity, GARCH models are nowadays widely used to represent the conditional volatility dynamics and to mimic other important stylized features of financial data.

Finally, a note on three important topics that are intrinsically related to temporal aggregation. First, since the seminal paper of Palm and Nijman (1984), a vast literature has analyzed the issue of *unobserved or missing* endogenous variables. Harvey (1981) proposes a clear definition of missing observation: a stock variable sampled less and less frequently with respect to its original model specification. The definition of missing observation, therefore, is extremely close to the definition of stock variable. For instance, moving from monthly to yearly frequency, yearly stock variables are observed, while monthly stock variables are not. Under a different perspective, it may be assumed that the monthly observations are missing, while the yearly ones are available. Then, it is necessary to compute the orders and the parameters of the disaggregate model (i.e. monthly) starting from the aggregate one (i.e. annual), estimated from the data. The whole subject is complicated by parameter identification issues that, quite expectedly, arise. In this paper we assume that the high frequency model is known and, on this basis, we infer the low frequency model. On the contrary, in a missing values context, the low frequency model is known but the high frequency model, the one that generates the observations, is unobservable. Therefore, while in this review we go from high to low frequency, the missing values approach goes from low to high frequency. For an exhaustive compendium of these issues we refer to the already cited Palm and Nijman (1984).

Second, temporal aggregation is not the only kind of aggregation. Many other economic phenomena may be analyzed from a cross-section perspective. For instance, the European Central Bank is potentially interested in measuring and forecasting economic variables for the Euro area. And this area is the sum, or the aggregation, of 15 countries. This scheme of aggregation, through individuals rather than through time, is called *contemporaneous aggregation*. It occurs when the observed series is the result of the sum of two or more individual series. As noted by Harvey (1981), if these individual series are known to follow stationary ARMA processes, it is possible to investigate whether the aggregate observed series follows an ARMA process, as well. The analogy between temporal aggregation and contemporaneous aggregation is evident. We refer the reader to Lütkepohl (1984 and 1986), Granger and Morris (1976) and Granger (1987, 1990), among others, for a thorough discussion.

Third, all the references quoted above assume that the high frequency model is either known or correctly specified. That is, there is no model uncertainty.[2] This immediately implies that the temporally aggregated model is also either known or correctly specified. Likewise, parameters are assumed to be exactly known and therefore there is no parameters uncertainty. However, as far as the temporal aggregation technique is concerned, parameters uncertainty is not an important issue as far as estimates are consistent and the sample size is large enough. And consistency is always fulfilled, using pseudo maximum likelihood arguments, as long as the model is either known or correctly specified.

To conclude, a comment on what we do not cover in this paper. Temporal aggregation of time series is a very vast field and mainly any subject in times series analysis may be investigated under this approach. In fact, the title of the article is very precise and already bounds the limits. Multivariate models, nonlinearities, long memory, random aggregation, time continuous or state space representations are issues that, though fascinating as they often show interesting features under temporal aggregation,[3] are not discussed and are left to future research.[4]

The paper is organized as follows. Section 2 introduces the notation and the aggregation schemes. Section 3 gives the intuitive foundations of the technique. Section 4 presents temporal aggregation for the easiest times series model, the AR. Section 5 discusses temporal aggregation for ARMA models. Section 6 is a summary of temporal aggregation for ARIMA models. Section 7 introduces seasonality in ARIMA models. Section 8 deals with temporal aggregation of a general ARIMAX model. Section 9 states the order conditions for temporal aggregation of GARCH models and presents some well known examples. Section 10 shows and empirical application based on macroeconomic data. Finally, our conclusions and suggestions for further research are drawn in Section 11.

2 NOTATION AND AGGREGATION SCHEMES

Let y_t, for $t = 0, 1, 2, \ldots$, be an *high frequency* or *disaggregate* time series observed at time t. Sample information for the *low frequency* or *aggregate* variable may be assumed to be available only every kth period $(k, 2k, 3k, \ldots)$, where k is an integer value larger that one. In general, we define the aggregate

[2] By 'model' we refer only the first two moments. Nowhere in the paper we assume some density for the error term.

[3] For instance, spurious causality in temporally aggregated multivariate models, i.e. causality of the low frequency multivariate GARCH process that is induced by temporal aggregation without any causal relationship at the high frequency (see Hafner, 2004) or the impact of temporal aggregation on causal relations in vector autoregressions (VARs) (see Breitung and Swanson, 2002).

[4] The book of Lütkepohl (1987) contains explanations and references concerning contemporaneous and temporal aggregation for vector ARMA processes (VARMA), while Hafner and Rombouts (2003) discuss the performance of various estimation techniques applied to temporally aggregated GARCH models. Random aggregation techniques are presented by Jorda and Marcellino (2004) and Aranzana and Veredas (2005). Nonlinearities issues are partly discussed in Proietti (2005) and Granger and Lee (1993). Links with time continuous models may be found in Nelson (1990) and Drost and Werker (1996). Finally, a study of the consequences of aggregation on long memory processes may be found in Granger (1980).

variable as

$$y_t^* = \sum_{j=0}^{A} w_j y_{t-j} = W(L) y_t. \tag{1}$$

This is a linear combination of current and past values of y_t, where $W(L) = \sum_{j=0}^{A} w_j L^j$ is a polynomial or order A in the lag operator L that determines the type of aggregation mechanism. The weights w_j, for $j = 0, 1, \ldots, A$, are exactly known.

Equation (1) embeds two important cases.

1. *Flow*: $A = k - 1$ and $w_j = 1$, for $j = 0, \ldots, A$, or $W(L) = 1 + L + \ldots + L^{k-1}$. The total flow since the last observation is measured, and hence the flow denomination. The integer k represents the temporal aggregation frequency or, equivalently, the order of aggregation. This means that aggregation of y_t variables is carried out over k periods. For instance, suppose that the observed time unit t is monthly and $k = 3$. Then we get as an outcome quarterly sums of the monthly realizations for y_t variable. In general, $y_t^* = \sum_{j=0}^{k-1} L^j y_t$. *Flow* scheme is also called *temporal aggregation*.
2. *Stock*: $A = 0$ and $w_0 = 1$. The pattern of observations is known as skipped observations. This time, the integer k represents the sampling frequency. If the observed time unit t is monthly and $k = 3$, the y_t variable is only observed every $3rd$ period. In general, $y_t^* = y_{kt}$. *Stock* scheme is also called *stock aggregation* or *systematic sampling*.

Other cases may be covered by (1). For example:

1. Weighted averaging: $w_j = \frac{\chi_j}{k}$, for $j = 0, \ldots, k - 1$. The weights are all different with predetermined values (χ_j), but they are all divided by k. Typically, the sum of all the weights is equal to one $\left(\sum_{j=0}^{k-1} \chi_j = 1 \right)$.
2. Phase averaging: $w_j = \frac{1}{k}$, for $j = 0, \ldots, k - 1$.

Flow and stock are the schemes more often found in economics. In the sequel, and for the sake of clarity, we shall focus on the flow case, although references to the stock case will be made whenever it helps to the understanding of the technique. Detailed results for the stock aggregation are relegated to the Appendix.

An important remark: note that the general scheme in (1) for flow, phase averaging and weighted averaging cases is a rolling sum. In other words, it is computed at every time t giving as result a sequence of sums that overlap for each of the $k - 1$ periods. However, the aggregate series is not overlapped. For instance, one year does not overlap with the next. To indicate the aggregate series, therefore, we introduce another time scale, T, that runs in kt periods. So that $t = 0, 1, 2, \ldots$, while $T = 0, k, 2k, \ldots$, as it is illustrated in Figure 1 for $k = 12$. Thus, we refer to the temporally aggregated series using the notation $y_T^* = y_{kt}^*$.

[FIGURE 1 ABOUT HERE]

Finally, we assume that the disaggregate series, y_t, follows a model of the type

$$\phi(L) y_t = \theta(L) \varepsilon_t, \tag{2}$$

where $t = 0, 1, 2, \ldots$, $\phi(L)$ and $\theta(L)$ are lag polynomials[5] and ε_t is an error term with possible time varying variance modelled via a GARCH model. Conversely, the temporally aggregated series, y_T^*, follows the model

$$\beta(B)y_T^* = \eta(B)\varepsilon_T^*, \tag{3}$$

where $T = 0, k, 2k, \ldots$, $\beta(B)$ and $\eta(B)$ are aggregate lag polynomials[6] and the operator B is in T time units, running in kt periods. The variable ε_T^* is an error term with possible time varying variance modelled via a GARCH model.

3 INTUITION I: ONLY WORDS

For the sake of clarity, we suppose that the process is stationary with constant variance. If the aggregate data (i.e. y_T^*) are a function (given by (1)) of the disaggregate data (i.e. y_t), we can think that the econometric model for y_T^*, given by (3), is also a function of the model for y_t, given by (2). This idea is represented for monthly and temporally aggregated annual data in Figure 2. The natural question is: how do these models link?

[FIGURE 2 ABOUT HERE]

In a univariate linear times series context, the expected value of y_t is a linear combination of past observations and past error terms. The number of lagged observations is given by the orders of the AR and MA polynomials that, in turn, are determined by the autocovariance structure of y_t. Once these orders are recovered, we estimate the parameters $(\hat{\phi}, \hat{\theta}) = (\hat{\phi}(y), \hat{\theta}(y))$.

The expected value of y_T^* is also a function of past values of y_T^* and past values of ε_T^*. However, what differs now with respect to y_t is that $y_T^* = y_{tk}^* = W(L)y_{tk}$. That is, the aggregate data are a function of the disaggregate data. Therefore the AR and MA aggregate polynomials orders are, through y_T^*, a function of the autocovariance structure of y_t. And the estimated parameters as well, $(\hat{\beta}, \hat{\eta}) = (\hat{\beta}(y), \hat{\eta}(y))$. Furthermore, following Fisher's paradigm, the estimated parameters should be such that they incorporate all the maximum information at the minimum cost. It means that all the information content in the high frequency sample, y_t, is in $(\hat{\phi}, \hat{\theta})$. And hence $(\hat{\beta}, \hat{\eta}) = (\hat{\beta}(\hat{\phi}, \hat{\theta}), \hat{\eta}(\hat{\phi}, \hat{\theta}))$: the parameters of the aggregate model are a function of the parameters of the disaggregate model.

Three conclusions may be already extracted from the previous paragraphs. First, we do not only aggregate data but we do also aggregate the model. In other words, the aggregate model is not estimated but *inferred* from the disaggregate model. Here 'inferred' has a twofold meaning: i) the lag structures of the AR and MA polynomials and ii) the corresponding parameters. And the method is sequential, first we derive the lag structures of the aggregate model and, subsequently, we infer its parameters.

Second, (ϕ, θ) are estimated with all the disaggregate observations. Thus, $(\hat{\beta}, \hat{\eta}) = (\hat{\beta}(\hat{\phi}, \hat{\theta}), \hat{\eta}(\hat{\phi}, \hat{\theta}))$ contains all the information of the high frequency sample. This definitely gives a more accurate estimate of the parameters, in terms of consistency and efficiency, than if they would be estimated from y_T^*, $(\hat{\beta}, \hat{\eta}) = (\hat{\beta}(y^*), \hat{\eta}(y^*))$, that has k times less observations. Moreover, and as we shall see, the function $(\hat{\beta}(\hat{\phi}, \hat{\theta}), \hat{\eta}(\hat{\phi}, \hat{\theta}))$ is deterministic, meaning that no loss of information is incurred or no extra risk is added.

Third, the use of this technique in practical applications implies that as soon as new disaggregate observations are available the aggregate parameters can be updated. This is a very useful tool for

[5] So far not necessarily stationary. They may include unit roots.

[6] So far not necessarily stationary. They may include unit roots.

situations where decisions are taken, say, annually, but information is available, for instance, monthly. No need to wait to the end of the year to re-estimate the model. Along the year the annual model may be updated as soon as monthly observations are released. And the updated model can be used for monitoring and forecasting.[7]

In an ARMA context, the two models are linked via a polynomial, that we denote by $T(L)$. This polynomial is a function of the roots of $\phi(L)$ and of the aggregation scheme in (1). This function drives us from one model to the other.

In general, the AR and MA polynomials of the disaggregate model expressed in terms of their roots are multiplied by $T(L)$: $T(L)\phi(L)y_t = T(L)\theta(L)\varepsilon_t$. The resulting AR polynomial, $T(L)\phi(L)$, has roots only divisible by $L^k = B$, the aggregate frequency. In this way y_t is transformed into y_T^*. Furthermore, i) the order of the AR polynomial remains the same under temporal aggregation and ii) since $T(L)$ is a function of the inverted roots of $\phi(L)$, the roots of $\beta(B)$ are the inverted roots of $\phi(L)$ powered by k, the temporal aggregation frequency. Multiplying by $T(L)$, in sum, the aggregate AR polynomial, its parameters and y_T^* appear naturally through a pure deterministic approach.

The aggregate AR part is the easiest one. The aggregate MA part, on the contrary, is more complicated. It is calculated multiplying the disaggregate MA polynomial by the $T(L)$ operator. Focusing on the RHS of (3), $T(L)\theta(L)\varepsilon_t$ aggregates to $\eta(B)\varepsilon_T^*$. The product $T(L)\theta(L)$ includes some AR components, the aggregation scheme (they are both present in $T(L)$) and the MA part. We therefore end up having two MA aggregate polynomials: i) the burdensome one, $T(L)\theta(L)\varepsilon_t$, resulting from the product of $\theta(L)\varepsilon_t$ by $T(L)$, and ii) the simple one, $\eta(B)\varepsilon_T^*$. Using some elementary deterministic rules, we may infer the order of the aggregate MA polynomial, $\eta(B)$, and its parameters. The trick consists in equating the autocovariance structures of both MA polynomials, $T(L)\phi(L)\varepsilon_t$ and $\eta(B)\varepsilon_T^*$. The result is a system of equations with as many equations as unknowns, that can be easily solved. This corresponds to a pure deterministic approach. Hence no extra risk is added and no information is lost.

When the process is not stationary, there is seasonality, exogenous variables or the variance of the error term has GARCH effects the technique becomes slightly more difficult. Nevertheless, the mechanism remains the same: there is always a polynomial function that links the two models. The aggregate AR polynomial (both in the mean and in the variance) is inferred straightforwardly. The MA structure is computed equating the autocovariance structures of the disaggregate and aggregate models.

4 INTUITION II:MORE THAN WORDS

In this section we outline the econometrics of temporal aggregation for the most basic linear time series models, namely pure autoregressive type models. The main reference is Amemiya and Wu (1972).

Let y_t follow an AR(p) model

$$\phi(L)y_t = \varepsilon_t, \tag{4}$$

where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ is an autoregressive polynomial of order p, $y_{t-i} = L^i y_t$ is the value of the y_t variable lagged i periods and ε_t is an independent white noise error term with mean equal zero and constant variance σ_ε^2. Let $\delta_j, j = 1, \ldots, p$, be the inverted roots of $\phi(L)$ polynomial. They are all assumed to lie inside the unit circle. Then, $\phi(L)$ may be factorized as $\phi(L) = \prod_{j=1}^{p}(1 - \delta_j L)$. Starting from this AR(p) model for y_t, the following proposition shows how to derive the appropriate specification for the temporally aggregated variable y_T^*.

[7] See Moulin *et al.* (2004) for an application to French public deficit.

Proposition 1 *The temporal aggregation of y_t as specified in model (4), denoted $\{y_T^*,\ T \in \mathbb{Z}\}$, may be represented by an ARMA(p,r) process where r, the maximum order of the aggregate moving average polynomial, is equal to*

$$r = \left\lfloor \frac{(p+1)(k-1)}{k} \right\rfloor, \tag{5}$$

with $\lfloor b \rfloor$ indicating the integer part of a real number b.

Proof. The proof fully stems from Amemiya and Wu (1972) and follows the intuitive guidelines outlined in Section 3. Indeed, the link between the models for y_t and y_T^* is given by the polynomial $T(L)$. For the AR(p) process, it takes the following form:[8]

$$T(L) = \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left[\frac{1-L^k}{1-L} \right]. \tag{6}$$

A closer look to $T(L)$ reveals that it has two parts. The first is a product of p terms, involving the inverted roots of $\phi(L)$. The second component involves the ratio of two polynomials that equals $\sum_{j=0}^{k-1} L^j$, that is the temporal aggregation scheme. In sum, the polynomial $T(L)$ has a component that transforms the roots of $\phi(L)$ into the roots of $\beta(B)$ and another component that accounts for the aggregation scheme.

Multiplying both sides of (4) by $T(L)$ we get

$$\prod_{j=1}^{p} \left[1-\delta_j^k L^k\right] y_t^* = T(L)\varepsilon_t, \tag{7}$$

where y_t^* is the temporally aggregated variable and with the temporal index t operating on the disaggregate time unit. Coming back to the ratio $\prod_{j=1}^{p} [(1-\delta_j^k L^k)(1-\delta_j L)^{-1}]$, its denominator contains the inverted roots of the AR polynomial and its numerator contains the same roots, but powered by the aggregation frequency. The powers of the product $T(L)\phi(L)$, consequently, are only divisible by the aggregation frequency. In other words, the only non-zero coefficients in $T(L)\phi(L)$ are those of powers of L divisible by k and the AR order is unchanged by temporal aggregation. The aggregate AR parameters are, simply, the disaggregate model inverted roots powered by the aggregation frequency

$$\beta(L)y_T^* = \prod_{j=1}^{p} \left[1-\delta_j^k L^k\right] y_T^*.$$

Following the only-words intuition, we now focus on the *RHS* of (7). It represents a moving average structure. Developing, we have

$$\begin{aligned} T(L)\varepsilon_t &= \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left(\sum_{i=0}^{k-1} L^i \right) \varepsilon_t \\ &= \prod_{j=1}^{p} \left(\sum_{i=0}^{k-1} \delta_j^i L^i \right) \left(\sum_{i=0}^{k-1} L^i \right) \varepsilon_t \\ &= \prod_{j=1}^{p} \left(\sum_{i=0}^{k-1} \delta_j^i L^i \right) \varepsilon_t^* = \prod_{j=0}^{p} \left(\sum_{i=0}^{k-1} \delta_j^i L^i \right) \varepsilon_t^*, \end{aligned}$$

[8]The general form of this polynomial will be given when we shall introduce seasonality. So far we start with the easiest form, adding terms as we add terms to the model.

with $\delta_0 = 0$. This corresponds to a linear combination of aggregate error terms. More precisely, $T(L)\varepsilon_t$ is a moving average of order $\lfloor (p+1)(k-1) \rfloor$. It is, however, expressed in terms of t rather than T. To switch the time frequency and to get the appropriate order of $\eta(B)$, we divide the order or $T(L)\theta(L)$ by k. That is, the order of $\eta(B)$ corresponds to $\lfloor k^{-1}(p+1)(k-1) \rfloor$. □

Next, we infer the parameters of the temporally aggregated model. For the AR part, they can be easily recovered as showed in (7). To compute the r parameters in $\eta(B)$, plus the variance $\sigma^2_{\varepsilon^*}$, we equate the autocovariance structures of $\eta(B)$ and $\prod_{j=0}^{p} \left(\sum_{i=0}^{k-1} \delta_j^i L^i \right) \varepsilon_t^*$. Since these are MA polynomials, only the variance and the r first autocovariances are different from zero. As the number of unknowns is also $r+1$, the problem consists in solving a system of equation with as many unknowns as equations.

Summarizing, an AR(p) process aggregates to an ARMA(p, r), with $r = \lfloor k^{-1}(p+1)(k-1) \rfloor$. The aggregate series follows an ARMA(p, r), that we can represent, letting $B = L^k$, as

$$\prod_{j=1}^{p} \left[1 - \delta_j^k B \right] y_T^* = (1 + \eta_1 B + ... + \eta_r B^r)\varepsilon_T^*, \tag{8}$$

or, more compactly, as

$$\beta(B) y_T^* = \eta(B) \varepsilon_T^*, \tag{9}$$

where ε_T^* process has zero mean and variance $\sigma^2_{\varepsilon^*}$.

In the remainder of the section we show, as an illustrative example, the easiest model that we can encounter in time series, an AR(1), with the lowest aggregation frequency that we may have, $k = 2$.

4.1 An example: Parameters of an aggregate AR(1) model

Consider the following AR(1) model

$$(1 - \phi L) y_t = \varepsilon_t, \tag{10}$$

where ε_t is an independent white noise error term with zero mean and constant variance σ^2_ε. Following the content of Proposition 1, the model for y_T^* is an ARMA(p, r), where $r = \lfloor k^{-1}(p+1)(k-1) \rfloor$. When $k = 2$, this is equivalent to an ARMA(1,1)

$$(1 - \beta B) y_T^* = (1 + \eta B) \varepsilon_T^*. \tag{11}$$

The $T(L)$ polynomial equals

$$T(L) = \left[\frac{1 - \phi^2 L^2}{1 - \phi L} \right] \left[\frac{1 - L^2}{1 - L} \right],$$

yielding the aggregate AR polynomial

$$T(L)(1 - \phi L) y_t = (1 - \phi^2 L^2) y_T^* = (1 - \beta B) y_T^*.$$

As a consequence, the aggregate AR parameter β is simply the disaggregate model inverted root powered by the aggregation frequency ($k = 2$), i.e. $\beta = \phi^2$. Hence, the AR order is unchanged by temporal aggregation.

Regarding the MA part, we have:

$$T(L)\varepsilon_t = (1 + \phi L)\varepsilon_t^* = (1 + \phi L)(1 + L)\varepsilon_t = \varepsilon_t + (1 + \phi)\varepsilon_{t-1} + \phi \varepsilon_{t-2}. \tag{12}$$

To compute the aggregate MA parameters η and $\sigma^2_{\varepsilon^*}$, we equate the autocovariance structures of $(1+\eta B)\varepsilon^*_T$ and (12). The variance of $(1+\eta B)\varepsilon^*_T$ is $\Gamma_0 = \left(1+\eta^2\right)\sigma^2_{\varepsilon^*}$ and the first-order autocovariance is $\Gamma_1 = \eta\sigma^2_{\varepsilon^*}$. Higher orders autocovariances are equal to zero. The variance of the model in (12) is $\gamma_0 = (1+(1+\phi)^2+\phi^2)\sigma^2_\varepsilon$, while the first-order autocovariance is $\gamma_1 = \phi\sigma^2_\varepsilon$. Higher orders autocovariances are null. This gives a system of two equations and two unknowns, $\Gamma_0 = \gamma_0$ and $\Gamma_1 = \gamma_1$, that is possible to solve for the aggregate parameters η and $\sigma^2_{\varepsilon^*}$

$$\Gamma_0 = \gamma_0 \Rightarrow \left(1+\eta^2\right)\sigma^2_{\varepsilon^*} = (1+(1+\phi)^2+\phi^2)\sigma^2_\varepsilon \Rightarrow \sigma^2_{\varepsilon^*} = \frac{(1+(1+\phi)^2+\phi^2)\sigma^2_\varepsilon}{(1+\eta^2)},$$

and

$$\Gamma_1 = \gamma_1 \Rightarrow \eta\sigma^2_{\varepsilon^*} = \phi\sigma^2_\varepsilon \Rightarrow \eta = \frac{\phi\sigma^2}{\sigma^2_{\varepsilon^*}} = \frac{\phi\sigma^2_\varepsilon}{\frac{(1+(1+\phi)^2+\phi^2)\sigma^2_\varepsilon}{(1+\eta^2)}} = (1+\eta^2)\frac{\phi}{(1+(1+\phi)^2+\phi^2)}$$
$$= (1+\eta^2)\rho_1,$$

where ρ_1 is the autocorrelation coefficient of order one. Then, the system reduces to the following second degree equation, of immediate solution:

$$\eta = (1+\eta^2)\rho_1 = 0 \Rightarrow \rho_1\eta^2 - \eta + \rho_1 = 0 \Rightarrow \eta^2 - \frac{\eta}{\rho_1} + 1 = 0.$$

In sum, the moving average aggregate parameters η and the aggregate variance $\sigma^2_{\varepsilon^*}$ are a function of the disaggregate AR inverted root ϕ and of the disaggregate variance σ^2_ε.

5 ARMA MODELS: TEMPORAL AGGREGATION

Let us introduce a more general class of model specifications: ARMA. The underlying process for y_t variable follows an ARMA(p, q) model

$$\phi(L)y_t = \theta(L)\varepsilon_t, \tag{13}$$

where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ and $\theta(L) = 1 + \theta_1 L + \ldots + \theta_q L^q$ are the autoregressive and moving average polynomials, of length p and q, respectively. The inverted roots of $\phi(L)$ and $\theta(L)$ are all inside the unit circle and the same polynomials have no roots in common. The variable ε_t is an independent white noise error term with mean equal to zero and constant variance σ^2_ε.

Up to now we have introduced the polynomial $T(L)$ as the cornerstone of the method. However, we have not justified it. Where does it come from? Why does it take the form that it takes and not another one? Why the aggregate AR parameters appear so easily while the MA are much more difficult to calculate? It is now time to answer to these questions. As a matter of fact, Brewer (1973) provides an interpretation of $T(L)$ polynomial that we present in the sequel of the section for flow aggregation, while a discussion on stock aggregation is deferred to Appendix A.1.

5.1 Aggregation of ARMA models: Order conditions

To obtain the aggregate model, each side of (13) has to be multiplied by $T(L)$. The only non-zero coefficients in the product $T(L)\phi(L)$ are those powers of L divisible by k. Consequently, the polynomial $T(L)$ must be such that the powers of the lag operator L appearing in the product $T(L)\phi(L)$ are all divisible by k. This requirement is expressed by Brewer in matrix form as:

$$T(L)\phi(L) = \begin{pmatrix} 1 & 0 & ... & 0 \\ t_1 & 1 & ... & 0 \\ t_2 & t_1 & 1 & 0 \\ ... & ... & ... & ... \\ t_p & t_{p-1} & ... & 1 \\ t_h & t_{h-1} & ... & t_{h-p} \\ 0 & t_h & ... & ... \\ 0 & 0 & t_h & ... \\ 0 & 0 & 0 & t_h \end{pmatrix} \begin{pmatrix} 1 \\ -\phi_1 \\ ... \\ -\phi_p \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ ... \\ \beta_1 \\ \beta_1 \\ ... \\ \beta_2 \\ \beta_2 \\ ... \\ \beta_s \end{pmatrix}. \tag{14}$$

A priori, the only known order is that of $\phi(L)$, $(p+1)\times 1$. The orders of $T(L)$ and $\beta(L)$ depend on h and s, respectively, that are unknown. We may however recover them as explained in what follows. The matrix corresponding to the $T(L)$ polynomial has size $(p+h+1)\times(p+1)$, while the matrix corresponding to the AR polynomial $\phi(L)$ has size $(p+1)\times 1$. Hence, the product $T(L)\phi(L)$ has dimension $(p+h+1)\times 1$. The RHS vector has size $(s+1)k\times 1$. A first restriction that helps us to recover h and s is that the number of rows on both sides of (14) is identical

$$p+h+1 = (s+1)k \Rightarrow p+h+1 = sk+k. \tag{15}$$

In addition, there are $t_1, t_2, \ldots, t_h$ unknown coefficients in the $T(L)$ polynomial to match with $(s+1)(k-1)$ equality conditions that have been imposed. On the RHS matrix, indeed, there are exactly $(s+1)(k-1)$ equal coefficients. Therefore a second restriction is

$$h = (s+1)(k-1) \Rightarrow h = sk+k-s-1. \tag{16}$$

Substituting (15) in (16), we get

$$h = p+h+1-s-1 \Rightarrow s = p. \tag{17}$$

Then, the order of the AR polynomial is unchanged after temporal aggregation ($p = s$).

Regarding the MA aggregate polynomial, $T(L)\theta(L)$, its order is $h+q$. Recall that $h = (p+1)(k-1)$. Hence $h+q = (p+1)(k-1)+q$. The autocovariance structure of order kr is $E[T(L)\theta(L)\varepsilon_t T(L)\theta(L)\varepsilon_{t-kr}]$, that is different from zero for $kr \leq (p+1)(k-1)+q$ or for $r \leq \lfloor k^{-1}\left((p+1)(k-1)+q\right)\rfloor$. Thus, the maximum aggregate MA order is $r = \lfloor k^{-1}\left((p+1)(k-1)+q\right)\rfloor$.

As example, consider the AR(1) model already introduced in subsection 4.1:

$$(1-\phi L)y_t = \varepsilon_t.$$

The temporal aggregation frequency, for simplicity, is still $k = 2$. In this case, the equation in matrix form (14) reduces to

$$\begin{pmatrix} 1 & 0 \\ t_1 & 1 \\ t_2 & t_1 \\ 0 & t_2 \end{pmatrix} \begin{pmatrix} 1 \\ -\phi \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ \beta_1 \\ \beta_1 \end{pmatrix},$$

which is equivalent to the following system of three equations in three unknowns, t_1, t_2 and β_1:

$$\begin{cases} t_1 - \phi = 1 \Rightarrow t_1 = 1+\phi \\ t_2 - t_1\phi = \beta_1 \\ -t_2\phi = \beta_1. \end{cases}$$

Substituting $t_1 = 1 + \phi$ inside $t_2 - t_1\phi = \beta_1$, we can solve the system by comparison:

$$\frac{\beta_1}{\phi} + \phi + \phi^2 = -\beta_1 \Rightarrow \beta_1 = \phi^2.$$

Thus, ϕ^2 is the value of the aggregate AR parameter, as we already know.

All the results so far obtained for temporal aggregation of ARMA(p,q) processes are summarized through the following proposition.

Proposition 2 *The temporal aggregation of y_t as specified in model (13), denoted $\{y_T^*,\ T \in \mathbb{Z}\}$, may be represented by an ARMA(p,r) process where r, the maximum order of the aggregate moving average polynomial, is*

$$r = \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor. \tag{18}$$

Proof. Multiplying both sides of (13) by the $T(L)$ operator, we have

$$\prod_{j=1}^{p} \left[1 - \delta_j^k L^k\right] y_t^* = T(L) \left(\sum_{j=0}^{q} \theta_j L^j \right) \varepsilon_t, \tag{19}$$

with $\theta_0 = 1$. The AR part of the model is treated exactly in the same way as in the pure autoregressive case. Regarding the disaggregate MA structure, the RHS of (19) may be expanded as

$$T(L) \left(\sum_{j=0}^{q} \theta_j L^j \right) \varepsilon_t = \left(\prod_{j=0}^{p} \sum_{i=0}^{k-1} \delta_j^i L^i \right) \left(\sum_{j=0}^{q} \theta_j L^j \right) \varepsilon_t.$$

Setting $\delta_0 = 1$, this is a MA structure of order $(p+1)(k-1)+q$. Thus, the aggregate series follows an ARMA(p,r) process, that we can represent, letting $B = L^k$, as

$$\prod_{j=1}^{p} \left[1 - \delta_j^k B\right] y_T^* = (1 + \eta_1 B + ... + \eta_r B^r)\varepsilon_T^*, \tag{20}$$

with

$$r = \lfloor k^{-1} \left((p+1)(k-1)+q\right) \rfloor.$$

□

5.2 An example: Parameters of an aggregate ARMA(1,1) model

As an illustration, let us focus on this disaggregate ARMA(1,1) model for the variable y_t

$$(1 - \phi L)y_t = (1 + \theta L)\varepsilon_t, \tag{21}$$

where ε_t is an independent white noise error term with mean equal zero and constant variance σ_ε^2.

Suppose now that the aggregation frequency is $k = 3$. This example is very likely to appear with real data if, for example, the disaggregate model has monthly frequency and we aggregate every three periods to get quarterly observations. Then the $T(L)$ operator becomes:

$$T(L) = \left[\frac{1 - \delta^3 L^3}{1 - \delta L}\right] \left(\sum_{j=0}^{2} L^j \right).$$

Multiplying the model in (21) by $T(L)$ we get:

$$\left[1-\delta^3L^3\right] y_t^* = \left[\frac{1-\delta^3L^3}{1-\delta L}\right](1+\theta L)\varepsilon_t^*. \tag{22}$$

The AR component is of order one and the aggregate AR parameter is equal to δ^3, i.e. the $\phi(L)$ inverted root powered by the aggregation frequency. For the MA part, we develop the RHS of (22) as follows, setting $\delta_0 = 1$

$$\begin{aligned}
&\left[\frac{1-\delta^3L^3}{1-\delta L}\right](1+\theta L)\varepsilon_t^* \\
&= \left(\sum_{j=0}^{2}\delta^j L^j\right)(1+\theta L)\left(\sum_{j=0}^{2}L^j\varepsilon_t\right) \\
&= \big(1+(1+\delta+\theta)L+(1+\delta+\delta^2+\theta+\theta\delta)L^2 \\
&+ (\delta+\delta^2+\theta+\theta\delta+\theta\delta^2)L^3+(\delta^2+\theta\delta+\theta\delta^2)L^4+\theta\delta^2L^5\big)\,\varepsilon_t,
\end{aligned} \tag{23}$$

which is a MA polynomial of order $(p+1)(k-1)+q=(1+1)(3-1)+1=5$. Therefore, the model in (21) aggregates to an ARMA$(1,r)$, where $r=\lfloor 3^{-1}(4+1)\rfloor = 1$.

Let now $B=L^k$ to operate on the aggregate time unit T. The temporal index $T=1,3,6,\dots$ is in quarterly frequency. Then, the aggregate series in (23) may be represented by the process

$$(1-\beta B)y_T^* = (1+\eta B)\varepsilon_T^*, \tag{24}$$

where $\beta=\delta^3$.

To compute the low frequency parameters η and $\sigma^2_{\varepsilon^*}$, we equate the variance-covariance matrices of ε_t and ε_T^* variables. For the disaggregate model in (23)

$$\begin{aligned}
\gamma_0 &= [1+(1+\delta+\theta)^2+(1+\delta+\delta^2+\theta+\theta\delta)^2+(\delta+\delta^2+\theta+\theta\delta+\theta\delta^2)^2 \\
&\quad +(\delta^2+\theta\delta+\theta\delta^2)^2+(\theta\delta^2)^2]\sigma^2_\varepsilon \\
\gamma_1 &= [(\delta+\delta^2+\theta+\theta\delta+\theta\delta^2)+(1+\delta+\theta)(\delta^2+\theta\delta+\theta\delta^2)+(1+\delta+\delta^2+\theta+\theta\delta)(\theta\delta^2)]\sigma^2_\varepsilon.
\end{aligned}$$

Note that, in real applications, the first step is to estimate the disaggregate ARMA model. Therefore, γ_0 and γ_1 are known constants as they depend on estimated parameters. For the temporally aggregated model in (24), $\Gamma_0=\left(1+\eta^2\right)\sigma^2_{\varepsilon^*}$ and $\Gamma_1=\eta\sigma^2_{\varepsilon^*}$. To compute the unknown parameters η and $\sigma^2_{\varepsilon^*}$, we must solve this elementary system in two equations in two unknowns

$$\begin{aligned}
&\Gamma_0=\gamma_0 \Rightarrow \left(1+\eta^2\right)\sigma^2_{\varepsilon^*}=\gamma_0 \Rightarrow \sigma^2_{\varepsilon^*}=\frac{\gamma_0}{(1+\eta^2)} \\
&\Gamma_1=\gamma_1 \Rightarrow \eta\sigma^2_{\varepsilon^*}=\gamma_1 \Rightarrow \eta=\frac{\gamma_1}{\sigma^2_{\varepsilon^*}}=\left(1+\eta^2\right)\frac{\gamma_1}{\gamma_0}=\left(1+\eta^2\right)\rho_1.
\end{aligned}$$

The system above can be therefore reduced to this second degree equation

$$\eta=\left(1+\eta^2\right)\rho_1 \Rightarrow \rho_1\eta^2-\eta+\rho_1=0 \Rightarrow \eta^2-\frac{\eta}{\rho_1}+1=0.$$

6 ARIMA MODELS: TEMPORAL AGGREGATION

In this section we make the further assumption that the data generating process is described by an ARIMA model,

$$\phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t, \tag{25}$$

where $\phi(L)$, $\theta(L)$ are stationary polynomials in the lag operator L of lengths p and q, respectively, ε_t is an independent white noise error term with zero mean and variance σ_ε^2, while d is a real integer denoting the order of integration. In the sequel we focus on flow aggregation. The stock case is deferred to Appendix A.2.

Temporal aggregation for ARIMA models is very similar to temporal aggregation for ARMA models. We only have to augment $T(L)$ to account for unit roots:

$$\bar{T}(L) = \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left[\frac{1-L^k}{1-L} \right]^{d+1}. \tag{26}$$

Notice the difference with $T(L)$ in (6), for the ARMA case. The second component is now powered by $d+1$ instead of 1, where d is the integration order. In fact, $(1-L^k)^d(1-L)^{-d}$ has a similar effect than the AR component. Multiplying the disaggregate model by $T(L)$ the high frequency unit roots vanish and only $(1-L^k)^d$ is left, i.e. a polynomial operating on T time units with d unit roots.

6.1 Aggregation of ARIMA models: Order conditions

We apply to model (25) a flow aggregation scheme. The proof stems from Weiss (1984). We do not show how to compute the aggregate parameters since this is equivalent to the ARMA case.

Proposition 3 *The temporal aggregation of y_t as specified in model (25), denoted $\{y_T^*,\ T \in \mathbb{Z}\}$, may be represented by an ARIMA(p,d,r) process where r, the maximum order of the aggregate moving average polynomial, is*

$$r = \left\lfloor \frac{p(k-1)+(d+1)(k-1)+q}{k} \right\rfloor. \tag{27}$$

Proof. Consider the process in (25). Let $\delta_1, \delta_2, \ldots, \delta_p$ be the reciprocals of the roots of $\phi(z)=0$ polynomial equation, assumed to lie inside the unit circle. We multiply each side of (25) by the augmented $\bar{T}(L)$ operator in (26), here below listed again,

$$\bar{T}(L) = \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left[\frac{1-L^k}{1-L} \right]^{d+1},$$

yielding

$$\begin{aligned} &\prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \prod_{j=1}^{p} [1-\delta_j L] \left[\frac{1-L^k}{1-L} \right]^{d+1} (1-L)^d y_t \\ &= \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left[\frac{1-L^k}{1-L} \right]^{d+1} \left(\sum_{j=0}^{q} \theta_j L^j \right) \varepsilon_t. \end{aligned}$$

Simplifying on the LHS

$$\prod_{j=1}^{p} [1-\delta_j^k L^k] \left(1-L^k\right)^d y_t^* = \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left[\frac{1-L^k}{1-L} \right]^{d} \left(\sum_{j=0}^{q} \theta_j L^j \right) \left(\sum_{j=0}^{k-1} L^j \right) \varepsilon_t.$$

The LHS is an ARI(p,d) operating in time T. Manipulating further the sums on the RHS we end up with

$$\prod_{j=1}^{p} [1-\delta_j^k L^k](1-L^k)^d y_t^* = \left(\sum_{l=0}^{p(k-1)+(d+1)(k-1)+q} \xi_l L^l \right) \varepsilon_t,$$

where ξ_l $(l = 0, \ldots, p(k-1) + (d+1)(k-1) + q)$ is a function of δ_j $(j = 1, \ldots, p)$ and θ_j $(j = 1, \ldots, q)$. The order of this latter polynomial in ε_t, on the RHS, is

$$\lfloor p(k-1) + (d+1)(k-1) + q \rfloor.$$

Thus, the aggregate series $\{y^*_T,\ T \in \mathbb{Z}\}$ follows an ARIMA(p, d, r) process, that we can represent, letting $B = L^k$, as

$$\prod_{j=1}^{p} \left[1 - \delta_j^k B\right](1 - B)^d y^*_T = (1 + \eta_1 B + ... + \eta_r B^r)\varepsilon^*_T, \tag{28}$$

with

$$r = k^{-1} \lfloor p(k-1) + (d+1)(k-1) + q \rfloor.$$

□

Note that the ε^*_T process is a white noise with zero mean and variance $\sigma^2_{\varepsilon^*}$. In addition, the moving average parameters η_j and the aggregate variance $\sigma^2_{\varepsilon^*}$ are functions of ϕ_j $(j = 1, \ldots, p)$, θ_j $(j = 1, \ldots, q)$ and σ^2_ε. The AR order is unchanged by temporal aggregation and the roots of the AR polynomial in the aggregate series are the kth powers of the inverted roots of the AR polynomial in the disaggregate series.

7 INTRODUCING SEASONALITY

Seasonality is present in many economic problems. One of the most well known examples are monthly economic indicators such as industrial production of unemployment. In this section we analyze the consequences of temporal aggregation in ARIMA models including seasonality.

Basically, it involves two issues: unit roots and dynamics. Seasonal movements may be non stationary and hence we may find unit roots at seasonal frequencies. What happens with the seasonal unit roots when we aggregate is the first question to answer. But seasonality may also be autocorrelated. We have, therefore, to analyze the consequences of temporal aggregation in the dynamics.

Temporal aggregation and seasonality are two terms that many times do not nicely harmonize. Some examples: suppose that we start with monthly observations that include annual seasonal pattern. That is, the time it takes to complete the seasonal cycle is one year. Suppose now that we compute quarters, i.e. we aggregate every 3 periods. The new seasonality is of 4 periods and is still annual, that is it takes one year to complete the seasonal cycle. But suppose now that we aggregate every 5 periods. Clearly, as the original seasonality is annual, there is still some seasonal cycle, but which one? What is the time it takes to complete the cycle? As we shall see, it takes 5 years. Therefore the aggregated seasonal cycle is different to the original cycle. Nevertheless, it is very rare in economics to find processes that aggregate to weird frequencies like in the example above. For instance, monthly observations almost always aggregate to quarters or one year, for which the seasonal cycle remains the same.

When aggregating seasonal cycles we may distinguish three cases: when the aggregation frequency is smaller than the cycle (e.g. we aggregate from monthly to quarterly), when it is the same (e.g. we aggregate from monthly to annual) and when it is larger (e.g we aggregate from monthly to biannual). In the first case, the temporally aggregated model still has some seasonal component. In the last two cases seasonality vanishes. But then, what happens to the dynamics and the to unit roots when seasonality vanishes? Everything becomes non seasonal. For instance, a seasonal AR(1) becomes a regular AR(1) and a seasonal unit root becomes a regular unit root.

To deal with seasonality, we introduce a new polynomial, $A(L)$, that has many similarities to $\bar{T}(L)$. In fact, $A(L)$ is to the seasonal part of the model as $\bar{T}(L)$ is to the non seasonal part. $A(L)$ is the link between the seasonality of the high and low frequency models. As in $\bar{T}(L)$, $A(L)$ has two components: one that deals with the dynamics and another that deals with the unit roots.

Due to the complex machinery that will be developed in the sequel of this section, we shall give several easy examples that will enhance the understanding. Last, and as usual, we shall discuss the method for flow aggregation. The stock aggregation case is deferred to Appendix A.3.

7.1 Intuition

Consider a general ARIMA$(p, d, q) \times (P, D, Q)_s$ model for the variable y_t

$$\phi(L)\Phi(L^s)(1-L)^d(1-L^s)^D y_t = \theta(L)\Theta(L^s)\varepsilon_t, \tag{29}$$

where ε_t is an independent white noise error term with zero mean, variance σ_ε^2. Standard regularity conditions are fulfilled by assumption. Let $\Phi(L^s) = 1 - \ldots - \Phi_P L^{Ps}$ and $\Theta(L^s) = 1 - \ldots - \Theta_Q L^{Qs}$ be polynomials in the seasonal lag operator of length Ps and Qs, respectively. In addition, let $\tau_1^s, \ldots, \tau_P^s$ be the reciprocals of the roots of the polynomial equation $\Phi(L^s)$, such that the factorization $\Phi(L^s) = \prod_{i=1}^{P} [1 - (\tau_i L)^s]$ holds. In the disaggregate series, we indicate with k the aggregation frequency, with s the seasonal frequency, with d the number of unit roots in levels and with D the number of unit roots in seasonality.

The difference with respect to the previous ARIMA model in (25) is the presence of seasonal polynomials. To account for this, we introduce a new operator, $A(L)$,

$$A(L) = \prod_{i=1}^{P} \left[\frac{1 - (\tau_i L)^{ks^*}}{1 - (\tau_i L)^s} \right] \left[\frac{1 - L^{ks^*}}{1 - L^s} \right]^D . \tag{30}$$

As $\bar{T}(L)$, it consists of two parts. The first includes the inverses of the roots of the seasonal autoregressive polynomial. The second account for the seasonal unit roots. However, notice that $A(L)$ does not include the aggregation scheme, while $\bar{T}(L)$ does. The reason is that instead than multiplying the high frequency model in (29) by $\bar{T}(L)$, we multiply it by $\bar{T}(L)A(L)$, and the aggregation scheme is already embedded in $\bar{T}(L)$.

Another difference with $\bar{T}(L)$ is given by s^*. It is the seasonal frequency of the temporally aggregated process. As already explained in the introduction, s^* may take weird values with lack of economic meaning. Indeed s^* depends on the original seasonal frequency, s, and on k. For instance, if the disaggregate process is monthly with annual frequency ($s = 12$) and we aggregate it every 3 periods (i.e. quarters), $s^* = 4$, which has a very clear economic meaning. However, if we aggregate every five periods, $s^* = 60$ or 5 years. That is, the temporally aggregated process takes five years to end up a cycle in the same month as the original process. The reason of this seasonal behaviour is that $k = 5$ is not a multiple of $s = 12$. It means that within a year (i.e. 12 months) there is not an exact number of aggregate periods, but it takes 5 year to have a 5 months period that ends up in a month multiple of 12.

Table 1 shows the cycle length of the temporally aggregated process when the original process is monthly with annual frequency, $s = 12$. The same weird cycle for $k = 5$ also appears for $k = 7, 8, 9, 10, 11$. The trick is that these numbers are not multiples of 12 and hence it takes a long time to complete a cycle at the same month as the monthly process.

[TABLE 1 ABOUT HERE]

Two interesting cases are when $k = 12$ and $k = 24$. For $k = 12$, the aggregation frequency equals the high frequency seasonal frequency, $k = s$. In this case seasonality vanishes. For $k = 24$ (or $k = 2s$), seasonality also vanishes. This drives us to some important conclusions: if $k < s$ there is still some seasonality in the temporally aggregated process. If k is multiple of s the seasonal cycle remains constant. If k is equal or larger than s seasonality vanishes.

7.2 Example I: No dynamics

The best way to grasp the meaning of these intuitions is to use a very simple model with no dynamics and only unit roots:

$$(1-L)^d(1-L^s)^D y_t = \varepsilon_t. \tag{31}$$

The polynomials $A(L)$ and $\bar{T}(L)$ are

$$\begin{aligned}\bar{T}(L) &= \left[\frac{1-L^k}{1-L}\right]^{d+1} \\ A(L) &= \left[\frac{1-L^{ks^*}}{1-L^s}\right]^{D}.\end{aligned}$$

We multiply (31) by the product $\bar{T}(L)A(L)$ and we only focus on the AR part of the model

$$\begin{aligned}&\left[\frac{1-L^k}{1-L}\right]^{d+1}\left[\frac{1-L^{ks^*}}{1-L^s}\right]^{D}(1-L)^d(1-L^s)^D y_t \\ &= (1-L^k)^d(1-L^{ks^*})^D y_t^* \\ &= (1-B)^d(1-B^{s^*})^D y_T^*,\end{aligned}$$

which is basically the same model as in (31) but in aggregate frequency, in the sense that it keeps the same structure, with the same number of regular and seasonal unit roots.

Following the last part of the previous subsection, notice what happens if k is a multiple of s. Then $s^* = s/k$ is an integer and

$$A(L) = \left[\frac{1-L^{ks^*}}{1-L^s}\right]^{D} = \left[\frac{1-L^{k\frac{s}{k}}}{1-L^s}\right]^{D} = 1.$$

Therefore, the polynomial that links the seasonality between the high and low frequency models becomes equal to 1. But because y_T^* has still some seasonal behaviour, we should have some seasonality in the aggregate model. Indeed, if $A(L) = 1$:

$$\begin{aligned}&\left[\frac{1-L^k}{1-L}\right]^{d+1}(1-L)^d(1-L^s)^D y_t \\ &= \left[\frac{1-L^k}{1-L}\right]^{d}\left[\frac{1-L^k}{1-L}\right](1-L)^d(1-L^s)^D y_t \\ &= (1-L^k)^d(1-L^s)^D y_t^* \\ &= (1-L^k)^d(1-L^{ks^*})^D y_t^* \\ &= (1-B)^d(1-B^{s^*})^D y_T^*.\end{aligned}$$

That is, we end up with the same model with no need of $A(L)$. Let us now give some numbers to k. Suppose that $k = 5$. The temporally aggregated model becomes $(1-L^5)^d(1-L^{5\times 12})^D y_T^* = (1-B)^d(1-B^{12})^D y_T^*$. Notice that $B^{12} = L^{5\times 12} = L^{60}$ and 60 is the number of periods that the aggregate series takes for completing the seasonal cycle at the same month as y_t. Suppose now that k is a multiple of s, say $k = 3$. The model becomes $(1-L^3)^d(1-L^{3\times 4})^D y_t^* = (1-B)^d(1-B^4)^D y_T^*$. Note that $B^4 = L^{12}$, that is the seasonal cycle of y_T^* is the same as the cycle of y_t.

Next, if k is larger or equal than s, we have seen that there is no seasonality left. Or $s^* = 1$. Then

$$\begin{aligned}
&\left[\frac{1-L^k}{1-L}\right]^{d+1}\left[\frac{1-L^k}{1-L^s}\right]^D (1-L)^d(1-L^s)^D y_t \\
&= (1-L^k)^d(1-L^k)^D y_t^* \\
&= (1-L^k)^{d+D} y_t^* \\
&= (1-B)^{d+D} y_T^*,
\end{aligned}$$

which is a model without seasonal unit roots, or, in other words, the seasonal unit roots in the disaggregate model are now regular unit roots. This is because when the aggregation frequency is a multiple of s, the seasonality at the low frequency vanishes. As a result, we still have the polynomial $A(L)$ but no seasonality in the aggregate series. Note that if $k = s$, k and s are both multiples of each other. It means that $A(L) = 1$ and we get the result as above.

7.3 Flow variables with seasonality: Order conditions

After this intuitive explanation of temporal aggregation and seasonal models, we report, in the following proposition, the orders of the aggregate AR and MA polynomials. The main reference is Weiss (1984).

Proposition 4 *The temporal aggregation of y_t as specified in model (29), denoted $\{y_T^*,\ T \in \mathbb{Z}\}$, may be represented by an $ARIMA(p,d,r)\times(P,D,R)_{s^*}$ where r, the maximum order of the regular moving average polynomial, is*

$$r = \left\lfloor \frac{(p+1)(k-1)+d(k-1)+q}{k} \right\rfloor, \tag{32}$$

while R, the maximum order of the seasonal moving average polynomial, is

$$R = \left\lfloor \frac{(P+D)s^*k+(Q-P-D)s}{k} \right\rfloor. \tag{33}$$

Proof. Multiplying both sides of (29) by $\bar{T}(L)A(L)$, we get:

$$\bar{T}(L)A(L)\phi(L)\Phi(L^s)(1-L)^d(1-L^s)^D y_t = \bar{T}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t.$$

Developing the LHS:

$$\begin{aligned}
&\prod_{j=1}^{p}\left[\frac{1-\delta_j^k L^k}{1-\delta_j L}\right]\left[\frac{1-L^k}{1-L}\right]^{d+1}\prod_{i=1}^{P}\left[\frac{1-(\tau_i L)^{ks^*}}{1-(\tau_i L)^s}\right]\left[\frac{1-L^{ks^*}}{1-L^s}\right]^D \\
&\times\prod_{i=1}^{p}[1-\delta_j L]\prod_{i=1}^{P}[1-(\tau_i L)^s](1-L)^d(1-L^s)^D y_t \\
&= \bar{T}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t.
\end{aligned}$$

Which implies

$$\begin{aligned}
&\prod_{j=1}^{p}[1-\delta_j^k L^k]\prod_{i=1}^{P}[1-(\tau_i L)^{ks^*}](1-L^k)^d(1-L^{ks^*})^D y_t^* \\
&= \bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t^*,
\end{aligned} \tag{34}$$

having defined

$$\bar{S}(L) = \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left[\frac{1-L^k}{1-L} \right]^d. \tag{35}$$

Note that $\bar{S}(L)$ is very similar to $\bar{T}(L)$ except for the aggregation scheme that is not present in $\bar{S}(L)$. On the LHS of (34) we have obtained an $ARI(p,d) \times (P,D)$ process. To calculate the order of the moving average part, $\bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t^*$, we develop the RHS of (34):

$$\begin{aligned}\bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t^* &= \prod_{j=1}^{p} \left[\frac{1-\delta_j^k L^k}{1-\delta_j L} \right] \left[\frac{1-L^k}{1-L} \right]^d \\ \times \prod_{i=1}^{P} \left[\frac{1-(\tau_i L)^{ks^*}}{1-(\tau_i L)^s} \right] & \left[\frac{1-L^{ks^*}}{1-L^s} \right]^D \theta(L)\Theta(L^s)\varepsilon_t^*. \end{aligned} \tag{36}$$

The order of the last polynomial in ε_t satisfies these equalities:

$$\begin{aligned} rk &= d(k-1) + p(k-1) + (k-1) + q, \\ Rk &= (P+D)s^*k + (Q-P-D)s. \end{aligned}$$

Thus, the temporally aggregated series $\{y_T^*,\ T \in \mathbb{Z}\}$ follows an ARIMA$(p,d,r) \times (P,D,R)_{s^*}$ process ($s^* = s/k$), that we can represent letting $B = L^k$ as:

$$\begin{aligned} &\prod_{j=1}^{p} \left[1-\delta_j^k B\right] \prod_{i=1}^{P} \left[1-\tau_i^{ks^*} B^{s^*}\right] (1-B)^d (1-B^{s^*})^D y_T^* \\ &= (1+\eta_1 B + ... + \eta_r B^r)(1+E_1 B + ... + E_R B^R)\varepsilon_T^*. \end{aligned} \tag{37}$$

□

Notice that, whenever $P = D = Q = 0$ (no seasonality), equation (32) becomes $r = k^{-1}\lfloor p(k-1) + q + (d+1)(k-1) \rfloor$, while R in (33) is null. Last, the derivations in this proof are slightly different from those that may be found in Weiss (1984). The main reason being that Weiss (1984) does not compute the maximum order R.

7.4 Example II: Some dynamics and aggregate parameters

Consider the well known Airline Model (Box, Jenkins, Reinsel, 1994)

$$(1-L)(1-L^{12})y_t = (1+\theta L)(1+\Theta L^{12})\varepsilon_t, \tag{38}$$

where ε_t is an independent white noise error process with zero mean and variance σ_ε^2. Within the same aggregation context, this model has been studied by Koreisha and Fang (2004). We also consider the same aggregation frequency, $k = 3$. The operators $\bar{T}(L)$ and $A(L)$ reduce to

$$\begin{aligned} \bar{T}(L) &= \left[\frac{1-L^3}{1-L} \right]^2 = \left[\frac{1-L^3}{1-L} \right] \left(\sum_{j=0}^{2} L^j \right) \\ A(L) &= 1, \end{aligned}$$

as k is a multiple of s. Multiplying the model in (38) by $\bar{T}(L)A(L)$, we get

$$(1-L^3)(1-L^{12})y_t^* = \left[\frac{1-L^3}{1-L} \right] (1+\theta L)(1+\Theta L^{12})\varepsilon_t^*, \tag{39}$$

with the usual notation $y_t^* = \left(\sum_{j=0}^{2} y_{t-j}\right)$ and $\varepsilon_t^* = \left(\sum_{j=0}^{2} \varepsilon_{t-j}\right)$. Equation (39) may be simplified in:

$$(1-L^3)(1-L^{12})y_t^* = \left(\sum_{i=0}^{2} L^i\right)(1+\theta L)(1+\Theta L^{12})\varepsilon_t^*. \tag{40}$$

And, using equations (32) and (33), the aggregated model for y_T^* corresponds to an ARIMA$(0,1,1) \times (0,1,1)_4$

$$(1-B)(1-B^4)y_T^* = (1+\eta_1 B)(1+E_1 B^4)\varepsilon_T^*. \tag{41}$$

We now compute the MA parameters (there are no AR parameters in this case). As usual, we calculate the non-zero autocovariances[9]

$$\begin{aligned}
\gamma_0 &= [19 + 32\theta + 19\theta^2 + 19\Theta^2 + 32\theta\Theta^2 + 19\theta^2\Theta^2]\sigma_\varepsilon^2 \\
\gamma_1 &= [4 + 11\theta + 4\theta^2 + 4\Theta^2 + 11\theta\Theta^2 + 4\theta^2\Theta^2]\sigma_\varepsilon^2 \\
\gamma_3 &= [4\Theta + 11\theta\Theta + 4\theta^2\Theta]\sigma_\varepsilon^2 \\
\gamma_4 &= [19\Theta + 32\theta\Theta + 19\theta^2\Theta]\sigma_\varepsilon^2.
\end{aligned}$$

And for the model in (41)[10]

$$\begin{aligned}
\Gamma_0 &= [1 + \eta_1^2 + E_1^2 + \eta_1^2 E_1^2]\sigma_{\varepsilon^*}^2 \\
\Gamma_1 &= [(1 + E_1^2)\eta_1]\sigma_{\varepsilon^*}^2 \\
\Gamma_3 &= \eta_1 E_1 \sigma_{\varepsilon^*}^2 \\
\Gamma_4 &= [(1 + \eta_1^2)E_1]\sigma_{\varepsilon^*}^2.
\end{aligned}$$

It is immediately evident that $\Gamma_4 = [(1+\eta_1^2)E_1]\sigma_{\varepsilon^*}^2 = E_1\sigma_{\varepsilon^*}^2 + \eta_1\Gamma_3$, hence Γ_3 is nested inside Γ_4 and should not be considered. As a consequence, the reduced system above is composed by three equations in three unknowns and it is exactly determined. Moreover, equating

$$\frac{\gamma_1}{\gamma_3} = \frac{1+\Theta^2}{\Theta} = \frac{\Gamma_1}{\Gamma_3} = \frac{1+E_1^2}{E_1} \Rightarrow E_1 = \Theta_1,$$

it comes out that $E_1 = \Theta$. This simplifies greatly the calculations. If E_1 is known, the variance of y_T^* in model (41) reduces to $\Gamma_0 = (1+\eta_1^2)\sigma_{\varepsilon^*}^2$. The first-order autocovariance is $\Gamma_1 = \eta_1\sigma_{\varepsilon^*}^2$. Higher orders autocovariances are all null. Coming to the model in (40), the variance is equal to

$$\gamma_0 \quad = \quad (19 + 32\theta + 19\theta^2)\sigma_\varepsilon^2,$$

while the first-order autocovariance is:

$$\gamma_1 \quad = \quad (4 + 11\theta + 4\theta^2)\sigma_\varepsilon^2.$$

Higher order autocovariances are all null. The unknown aggregate parameters η_1 and $\sigma_{\varepsilon^*}^2$ may be determined equating the variance-covariance structures of the models in (41) and in (40), as follows:

$$\begin{aligned}
(19\theta^2 + 32\theta + 19)\sigma_\varepsilon^2 &= (1+\eta_1^2)\sigma_{\varepsilon^*}^2 \\
(4\theta^2 + 11\theta + 4)\sigma_\varepsilon^2 &= \eta_1\sigma_{\varepsilon^*}^2.
\end{aligned} \tag{42}$$

Hence, given the monthly model in (38) and its corresponding estimated parameters θ, Θ and σ_ε^2, it is possible to solve the system in (42) to calculate the aggregate parameters of model (41), i.e. η_1 and $\sigma_{\varepsilon^*}^2$. In addition, the equality $E_1 = \Theta$ holds.

[9]The autocovariance of order 5 is different from zero but equal to that of order 3. It is therefore redundant and may be eliminated from the system.

[10]As before, the autocovariance of order 5 is different from zero but equal to that of order 3.

8 TEMPORAL AGGREGATION OF ARIMAX MODELS

It happens very often in econometric models that an economic variable is not only explained by itself but also by other variables, assumed to be exogenous and otherwise called explanatory (generally, they are denoted by the letter x). In this section we study temporal aggregation when these explanatory variables are present. For the sake of simplicity, we assume that only one exogenous variable is introduced in the model and that there is no seasonality. Extensions to more than one variable are straightforward.

Four important clarifications before starting. First, at the beginning of the article we made the distinction between different types of aggregation schemes, being stock and flow the most common in economics. When the model includes explanatory variables, it may happen that its aggregation scheme differs from the aggregation scheme of y_t. Many possibilities are at hand but in order to introduce some variety we assume that y is flow, as usual, while x is stock. Second, the new twist in complexity comes by the fact that the explanatory variable, although exogenous, has its own ARIMA model. This means that when we aggregate the model for y, the orders of the temporally aggregated model are going to be a function of the orders of the model for x. But this also implies that we are going to get not only one equation that defines the order of the aggregate MA polynomial (as we have done up to now), but also one for the aggregate component of x. Third, we assume that x is observed at the high frequency. This is an assumption that may not be fulfilled in real applications. For instance, if y is the monthly industrial production series and x the GDP, this requirement is not met, since GDP is not observed monthly. Nevertheless, there are plenty of economic applications where both y and x are observed at all the frequencies. Last, Brewer (1973) carries out temporal aggregation for ARMAX models, but not for integrated processes. Therefore, the proof in this section may be considered as novelty.

We consider an ARIMAX(p, d, q) model with complete lag structure in the errors and in the *stock* exogenous variable:

$$\phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t + C(L)(1-L)^{\tilde{d}} x_t, \tag{43}$$

where $\phi(L)$ and $\theta(L)$ are the usual polynomials in the lag operator L of length p and q, respectively, and ε_t is an independent white noise error term with mean zero and constant variance σ_ε^2. Standard regularity conditions are fulfilled. Let x_t be a stock exogenous variable and let $C(L)$ be its associated polynomial of length m. The underlying assumption is that x_t follows an ARIMA$(v, \tilde{d}, w)$ process

$$D(L)(1-L)^{\tilde{d}} x_t = F(L)u_t, \tag{44}$$

where $D(L)$ and $F(L)$ are polynomials of length v and w, respectively, $\tilde{d}$ is the integration order and u_t has mean zero, variance σ_u^2 and is independent by ε_t.

The problem consists in performing temporal aggregation of the ARIMAX model in (43) with a stock exogenous variable obeying (44). The following proposition gives the order conditions and the number of past lags of x in the temporally aggregated model. The proof is slightly different from the previous ones because of the presence of x. Indeed, due to the temporal aggregation mechanism, some of the x are unobserved in the temporally aggregated specification. This problem is overcome expressing the unobserved x terms as a function of their own past, using their own model (44). This is done, concretely, applying to (44) an operator that shares many similarities with $T(L)$.

Proposition 5 *With the ARIMAX(p,d,q) model specified in (43) and with the stock exogenous variable following (44), the temporally aggregated series of y_t, denoted $\{y_T^*,\ T \in \mathbb{Z}\}$, may be represented by an ARIMAX(p,d,r) where r, the maximum aggregate moving average order, is the greater of*

$$\left\lfloor k^{-1}[(p+d+1)(k-1)+q] \right\rfloor \ \textit{and}$$
$$\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}] \right\rfloor + \left\lfloor k^{-1}[(v+\tilde{d})(k-1)+w] \right\rfloor, \tag{45}$$

while the maximum order of the independent stock exogenous variable is

$$\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}]\right\rfloor + v + \tilde{d}. \tag{46}$$

Proof. First, we apply the temporal aggregation operator $\bar{T}(L)$ in (26) to both sides of (43), obtaining

$$\bar{T}(L)\phi(L)(1-L)^d y_t = \bar{T}(L)\theta(L)\varepsilon_t + \bar{T}(L)C(L)(1-L)^{\tilde{d}}x_t. \tag{47}$$

Equation (47) may be rewritten as

$$\prod_{j=1}^{p}\left[1-\delta_j^k L^k\right]\left(1-L^k\right)^d y_t^* = \bar{S}(L)\theta(L)\varepsilon_t^* + \bar{T}(L)C(L)(1-L)^{\tilde{d}}x_t. \tag{48}$$

The treatment for the AR part is identical to the pure ARIMA model (see Proposition 3). In particular, the aggregate autoregressive order remains equal to p.

Let us therefore focus on the stock exogenous variable, i.e. on the $\bar{T}(L)C(L)(1-L)^{\tilde{d}}x_t$ term, whose polynomial length is $[(p+d+1)(k-1)+m+\tilde{d}] = [(p+d+1)k+(m+\tilde{d}-p-d-1)] = [(p+d+1)k+j] = [ik+j]$, with $j = (m+\tilde{d}-p-d-1)$ and i an integer number. Hence, in (48), there are terms such as $(\ldots)x_{t-ik-j}$ (with $0 < j < k$), i.e. some of the x_t are unobserved in the temporally aggregated specification. In other words, because of the different lag lengths of $C(L)$ and $\phi(L)$, there are lags of x in the temporally aggregated model that are not divisible by the aggregation frequency, i.e. they are in between $t-k$ and $t-(k-1)$. And these terms are given by the *temporal location of the unobserved values* (Brewer, 1973), i.e. they are a function of j.

Model (44) has to be used to express these unobserved terms as a function of the observed terms of the same variable and random innovations u_t (the errors of the model in x), that can be absorbed inside the MA part of the temporally aggregated model. To do this, we filter both sides of (44) with the operator $\bar{P}(L) = P(L)\left[\frac{1-L^k}{1-L}\right]^{\tilde{d}}$, of generic length $g+\tilde{d}(k-1)$, obtaining

$$P(L)\left[\frac{1-L^k}{1-L}\right]^{\tilde{d}} D(L)(1-L)^{\tilde{d}}x_t = P(L)\left[\frac{1-L^k}{1-L}\right]^{\tilde{d}} F(L)u_t$$

$$\Rightarrow P(L)D(L)(1-L^k)^{\tilde{d}}x_t = P(L)\left[\frac{1-L^k}{1-L}\right]^{\tilde{d}} F(L)u_t. \tag{49}$$

The only non-zero coefficients in the product $P(L)D(L)$ on the LHS of (49) are those corresponding to L^{k-j}, L^{2k-j}, $\ldots$, L^{sk-j} lag operators: indeed, i) the stock scheme is applied to the exogenous variable x_t (i.e. L^{sk}) and ii) some of the x_t are unobserved in the temporally aggregated specification (i.e. L^{-j}).

This requirement is expressed by Brewer (1973) in matrix form as:

$$\begin{pmatrix} 1 & 0 & ... & 0 \\ p_1 & 1 & ... & 0 \\ p_2 & p_1 & 1 & 0 \\ ... & ... & ... & ... \\ p_v & ... & ... & 1 \\ p_g & ... & ... & ... \\ 0 & p_g & ... & ... \\ 0 & 0 & p_g & ... \\ 0 & 0 & 0 & p_g \end{pmatrix} \begin{pmatrix} 1 \\ -d_1 \\ ... \\ -d_v \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ ... \\ 0 \\ d^*_{k-j} \\ 0 \\ ... \\ 0 \\ d^*_{2k-j} \\ 0 \\ ... \\ 0 \\ d^*_{sk-j} \end{pmatrix}. \tag{50}$$

To recover the unknown orders g (i.e. generic length of the operator $P(L)$) and s (i.e. AR polynomial length in the aggregate series) as a function of k (i.e. aggregation frequency) and v (i.e. AR polynomial length in the disaggregate series), we proceed equating the number of rows on both sides of (50)[11]

$$1 + v + g = 1 + sk - j, \tag{51}$$

and equating the number of unknown coefficients in $P(L)$ with the $s(k-1) - j$ equality restrictions that have been implicitly imposed on the RHS vector of (50):

$$g = s(k-1) - j. \tag{52}$$

Solving this system of two equations brings to:

$$\begin{aligned} s &= v \\ g &= v(k-1) - j. \end{aligned}$$

Therefore, the order of the aggregate AR polynomial for the stock exogenous variable $(1-L)^{\tilde{d}}x_t$ is unchanged after temporal aggregation (i.e. $s = v$). In addition, the generic length of the $P(L)$ polynomial used to filter equation (49), $g = v(k-1) - j$, depends by the *temporal location of the unobserved value*, i.e. it is a function of j.

Once the length of the polynomial $P(L)$ is determined, we come back to $P(L)D(L)(1-L^k)^{\tilde{d}}x_t$, on the LHS of (49). The earliest observed value of the independent variable $(1-L^k)^{\tilde{d}}x_t$ is

$$(1-L^k)^{\tilde{d}}x_{t-ik-j-g-v} = (1-L^k)^{\tilde{d}}x_{t-ik-j-v(k-1)+j-v} = (1-L^k)^{\tilde{d}}x_{t-(i+v)k}, \tag{53}$$

a quantity which is independent by the temporal location of the unobserved value, i.e. it is not a function of j. This means that, after filtering (49) with the operator $\bar{P}(L)$, we are able to express the unobserved value of $(1-L^k)^{\tilde{d}}x_{t-ik-j}$ as a function of $(1-L^k)^{\tilde{d}}x_{t-(i+1)k}$, $(1-L^k)^{\tilde{d}}x_{t-(i+2)k}$, ... Equation (53) is still valid if we replace the generic integer i with $\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}]\right\rfloor$, which is the length of $\bar{T}(L)C(L)(1-L)^{\tilde{d}}x_t$ term in (48):

$$(1-L^k)^{\tilde{d}}x_{t-\left\lfloor k^{-1}[(p+d+1)(k-1)+m+\tilde{d}]\right\rfloor k-vk}.$$

[11] We are following an identical procedure to the one in Section 5.1.

As a direct consequence, the earliest observed value of the independent variable x_t is

$$x_{t-\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}]\right\rfloor k-vk-\tilde{d}k}, \tag{54}$$

and the maximum lag on the independent variable is $\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}]\right\rfloor + v + \tilde{d}$, as indicated in (46).

Finally, we come back to the MA part, i.e. $P(L)\left[\frac{1-L^k}{1-L}\right]^{\tilde{d}} F(L)u_t$, on the RHS of (49). Its polynomial order is $g + \tilde{d}(k-1) + w = v(k-1) - j + \tilde{d}(k-1) + w = (v+\tilde{d})(k-1) + w - j$. The earliest relevant u-value corresponding to $(1-L^k)^{\tilde{d}} x_{t-ik-vk}$ is:

$$u_{t-ik-j-g-w} = u_{t-ik-j-(v+\tilde{d})(k-1)+j-w} = u_{t-ik-[(v+\tilde{d})(k-1)+w]}.$$

Replacing the generic integer i with $\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}]\right\rfloor$, we end up with:

$$u_{t-ik-[(v+\tilde{d})(k-1)+w]} = u_{t-\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}]\right\rfloor k-[(v+\tilde{d})(k-1)+w]}. \tag{55}$$

Therefore, we conclude that the moving average order is the larger between $\left\lfloor k^{-1}[(p+d+1)(k-1)+q]\right\rfloor$ and $\left\lfloor k^{-1}[(p+d+1)(k-1)+m-1+\tilde{d}]\right\rfloor + \left\lfloor k^{-1}[(v+\tilde{d})(k-1)+w]\right\rfloor$, as reported in (45). □

9 TEMPORAL AGGREGATION OF GARCH MODELS

In this section we review the main results on temporal aggregation of univariate ARMA-GARCH models, that have been widely used to explain conditional heteroskedasticity of financial time series. As the main subject of the section is the time-varying variance, we rely on a simple ARMA model although extensions to ARIMA models with seasonality and/or exogenous variables are straightforward.

The issue of temporal aggregation in financial time series is of great importance. In other fields of economics, practice and theory tell us about the optimal sampling frequency. Months, quarters or years seem to be the most used ones. However, in financial markets such an a priori optimal frequency does not exist. Investment bankers may argue that the optimal frequency is a month, if not a quarter, mutual fund managers are ultimately interested in weekly data and hedge fund managers think in daily, if not intra-daily, basis.

An ARMA(p,q)-GARCH(P,Q) model is defined as

$$\begin{aligned} \phi(L)y_t &= \theta(L)\varepsilon_t, \ \varepsilon_t \sim D(0,h_t) \\ h_t &= \psi + \sum_{i=1}^{P} b_i h_{t-i} + \sum_{i=1}^{Q} a_i \varepsilon_{t-i}^2. \end{aligned} \tag{56}$$

where $\{\varepsilon_t, t \in \mathbb{Z}\}$ is a sequence of stationary errors with finite fourth moments and $\phi(L)$ and $\theta(L)$ are stationary autoregressive and moving average polynomials of length p and q, respectively. Stationarity condition for the GARCH model is $\sum_{i=1}^{P} b_i + \sum_{i=1}^{Q} a_i < 1$. The GARCH model can be rewritten as

$$b(L)h_t = \psi + [a(L)-1]\varepsilon_t^2, \tag{57}$$

where $a(L) = 1 + \sum_{i=1}^{Q} a_i L^i$ and $b(L) = 1 - \sum_{i=1}^{P} b_i L^i$. The polynomials $b(L)$ and $b(L)+1-a(L)$ are assumed to be invertible and to have roots outside the unit circle.

As a difference with the previous sections, a new concept has to be introduced here: closeness.[12] Closeness is a broad term in econometrics. In the aggregation setting it means that when aggregating the model remains the same. But what does *the model remains the same* mean? Does it mean that the distribution remain the same? Or does it mean that the lag orders remain the same? Or, furthermore, does it mean that errors remain uncorrelated under aggregation? Drost and Nijman (1993) introduce three definitions of GARCH models (strong, semi-strong and weak) and show that only one (weak) remains closed under temporal aggregation. In a nutshell, strong GARCH means that errors, standardized by the conditional standard deviation, are *i.i.d* with zero mean and unit variance. Semi-strong GARCH goes down to the first two moments, and it defines a GARCH model as a martingale difference sequence, i.e. conditional mean equal to zero and variance equal to the GARCH model. Weak GARCH goes even further down in the requirements and it needs only that differences between expected and realized first and second moments of the residuals are uncorrelated.[13]

Drost and Nijman (1993) show that only weak GARCH models are closed under temporal aggregation. That is, when going from y_t to y_T^* the distribution may change or ε_T^* may not be a martingale difference sequence anymore but differences between realized and expected values are uncorrelated. This idea is defined in what follows.

Definition 1 *We define the sequence $\{\varepsilon_t, t \in \mathbb{Z}\}$ to be generated by a symmetric* ***weak*** *GARCH error process if ε_t variables are uncorrelated and if ψ, $a(L)$ and $b(L)$ can be chosen such that*

$$\begin{aligned} P[\varepsilon_t|\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] &= 0 \\ P[\varepsilon_t^2|\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots] &= h_t, \end{aligned} \tag{58}$$

where $P[x_t|\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots]$ represents the best linear predictor in terms of 1, ε_{t-1}, ε_{t-2}, $\ldots$, ε_{t-1}^2, ε_{t-2}^2, $\ldots$. That is,

$$E\left(x_t - P[x_t|\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots]\right)\varepsilon_{t-i}^r = 0 \ for \ i \geq 1 \ \ and \ r = 0, 1, 2.$$

Therefore, a GARCH model is weak if the expected difference between the realized and the projected value is orthogonal, up to the second moment, with the realized value.

We now show and proof that ARMA-GARCH models with weak GARCH errors are closed under temporal aggregation in the sense defined earlier. The proof is structured in two steps. The first step is as usual, inferring the aggregate conditional mean (i.e. the ARMA part of the model) from the high frequency one. The second step deals with the variance. The GARCH model is defined on the variance of the residuals and hence treatment is slightly different from the conditional mean - though it resembles. The proof for the second moment starts rewriting the GARCH models as an ARMA model. Then, following the logic of the previous sections, an operator is applied to the transformed model. However, this operator differs from all the others. It differs because the GARCH model is not on y_T^* but on ε_T^*. We present the order conditions for the flow case, relegating the stock case to Appendix A.4.

Next proposition deals with the orders of the aggregate GARCH model, nothing being said about the parameters. Subsection 9.2 explains how to obtain the parameters in the case of stock aggregation for a GARCH(1,1) with $k = 2$. It actually replicates Example 1 of Drost and Nijman (1993).

[12]This sentence lacks of rigor but we hope that the reader will forgive us. Closeness is not a concept related to the aggregation of GARCH models but much more general, that can also be applied to ARIMA models. However, for the ease of the exposition, and because it was not strictly needed, we preferred not to introduce it earlier.

[13]This third definition will become clearer shortly.

Proposition 6 *The temporal aggregation of y_t as specified in (56), denoted $\{y_T^*,\ T \in \mathbb{Z}\}$, may be represented by an ARMA(p,r) with weak GARCH(R,R) errors where*

$$
\begin{aligned}
r &= \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor \\
R &= \tilde{r} + \frac{1}{2} r(r+1),
\end{aligned}
$$

and with $\tilde{r} = \max(P, Q)$.

Proof. In the first step, the ARMA part fully mimics the proof for ARMA with constant variance. That is, we multiply $\phi(L)y_t = \theta(L)\varepsilon_t$ by

$$
T(L) = \prod_{i=1}^{p} \left[\frac{1-\delta_i^k L^k}{1-\delta_i L} \right] \left[\frac{1-L^k}{1-L} \right],
$$

where, as usual, δ_i, $i = 1, \ldots, p$ represent the inverted roots of $\phi(L)$ polynomial, assumed to be stationary. This gives

$$
\prod_{i=1}^{p} (1-\delta_i^k L^k) y_t^* = T(L)\theta(L)\varepsilon_t.
$$

And we know that this model corresponds to an ARMA(p,r) in aggregate time units,

$$
\beta(B) y_T^* = \eta(B) \varepsilon_T^*,
$$

with $r = \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor$.

In the second step, we infer the GARCH structure for ε_T^* taking into account that it actually derives from $T(L)\theta(L)\varepsilon_t$. To do this, we first express the GARCH model (57) as an ARMA. Let $\iota_t = \varepsilon_t^2 - h_t$. Substituting in (57) we get $(1 - a(L) + b(L))\varepsilon_t^2 = \psi + b(L)\iota_t$. Or, equivalently,

$$
C(L)\varepsilon_t^2 = \psi + b(L)\iota_t, \tag{59}
$$

where $C(L) = 1 - a(L) + b(L)$ is a polynomial of order $\tilde{r} = \max(P, Q)$ and $b(L)$ is of order P. Moreover, we may write $C(L)$ in terms of its inverted roots, i.e. $C(L) = \prod_{i=1}^{\tilde{r}} (1 - \zeta_i L)$. This is the GARCH model for the high frequency data. We are interested in the corresponding GARCH model for the low frequency data.

To link both models we use two polynomials that share certain similarities with $T(L)$. The first one,

$$
\hat{T}(L) = \prod_{i=1}^{\tilde{r}} \left[\frac{1-\zeta_i^k L^k}{1-\zeta_i L} \right],
$$

simply allows to recover the aggregate AR parameters in (59). The covariance structures of $\eta(B)\varepsilon_T^*$ and $T(L)\theta(L)\varepsilon_t$ are the same and, as a consequence, $\eta(L^k)\varepsilon_T^* = T(L)\theta(L)\varepsilon_t$ or $\varepsilon_T^* = \eta(L^k)^{-1}T(L)\theta(L)\varepsilon_t$, for all t multiple of k. In addition, note that $\varepsilon_T^* = \eta(L^k)^{-1}T(L)\theta(L)\varepsilon_t = \Psi(L)\varepsilon_t$ is an uncorrelated sequence where $\Psi(L) = \eta(L^k)^{-1}T(L)\theta(L)$ is a polynomial of infinite order. But, obviously, we need a finite polynomial structure for ε_t. Therefore, as a second polynomial, we define the finite sum of squares

$$
\Psi^2(L) = \sum_{i=0}^{k-1} \psi_i^2 L^i.
$$

Then we multiply (59) by the product of $\hat{T}(L)$ times $\Psi^2(L)$. This product of polynomials constitutes the link between the two error terms and hence links the GARCH models for ε_t and ε_T^*, delivering

$$\prod_{i=1}^{\tilde{r}} \left[1-\zeta_i^k L^k\right] \Psi^2(L)\varepsilon_t^2 = \hat{T}(1)\Psi^2(1)\psi + \hat{T}(L)b(L)\Psi^2(L)\iota_t. \tag{60}$$

Since $(\varepsilon_T^*)^2 = (\Psi(L)\varepsilon_t)^2$ for all t multiple of k, (60) can be rewritten in terms of $(\varepsilon_T^*)^2$,

$$\prod_{i=1}^{\tilde{r}} \left[1-\zeta_i^k L^k\right] (\varepsilon_T^*)^2 = \hat{T}(1)\Psi^2(1)\psi + \hat{T}(L)b(L)\Psi^2(L)\iota_t + \prod_{i=1}^{\tilde{r}} \left[1-\zeta_i^k L^k\right] \left((\Psi(L)\varepsilon_t)^2 - \Psi^2(L)\varepsilon_t^2\right). \tag{61}$$

In (61) we already have a GARCH structure for the low frequency model. The LHS is an AR polynomial of order $\tilde{r}$, while the second and third terms on the RHS determine the MA part. We remark that $\left((\Psi(L)\varepsilon_t)^2 - \Psi^2(L)\varepsilon_t^2\right)$ has only $\varepsilon_{t-i}\varepsilon_{t-j}, i \neq j$ terms, as the ε_t^2 in $(\Psi(L)\varepsilon_t)^2$ vanish with those of $\Psi^2(L)\varepsilon_t^2$. Hence the second and third terms on the RHS are uncorrelated, since η_t and $\varepsilon_{t-i}\varepsilon_{t-j}$ are uncorrelated by assumption ($\forall i \neq j \in \mathbb{Z}$). We now have to find the order of this MA structure. However, due to the presence of $(\Psi(L)\varepsilon_t)^2$, the MA structure on the RHS is not finite. To obtain a finite structure it is necessary to multiply (61) by another polynomial.

The polynomial $\eta(L^k)\Psi(L)\varepsilon_t = \eta(L^k)\eta(L^k)^{-1}T(L)\theta(L)\varepsilon_t = T(L)\theta(L)\varepsilon_t$ is finite and of order $(p+1)(k-1)+q$. Therefore

$$\sum_{j=0}^{r} \eta_j \psi_{i-jk} = 0, \;\; for\ i > (p+1)(k-1)+q. \tag{62}$$

This difference equation in ψ of order $r = \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor$ determines a difference equation in ψ^2 of order $\frac{1}{2}r(r+1)$, of the type

$$\sum_{j=0}^{\frac{1}{2}r(r+1)} \tilde{\psi}_j \psi_{i-jk}^2 = 0, \;\; for\ i > (p+1)(k-1)+q+\frac{1}{2}kr(r-1).$$

Defining $\tilde{\Psi}(L^k) = \sum_{i=0}^{\frac{1}{2}r(r+1)} \tilde{\psi}_i L^{ik}$, we can observe that $\tilde{\Psi}(L^k)\Psi^2(L)$ is a polynomial of *finite* order $(p+1)(k-1)+q+\frac{1}{2}kr(r-1)$.

Multiplying (61) by $\tilde{\Psi}(L^k)$ yields

$$\tilde{\Psi}(L^k)\prod_{i=1}^{\tilde{r}} \left[1-\zeta_i^k L^k\right] (\varepsilon_T^*)^2 = \tilde{\Psi}(1)\hat{T}(1)\Psi^2(1)\psi +$$
$$\tilde{\Psi}(L^k)\hat{T}(L)b(L)\Psi^2(L)\iota_t + \tilde{\Psi}(L^k)\prod_{i=1}^{\tilde{r}} \left[1-\zeta_i^k L^k\right] \left((\Psi(L)\varepsilon_t)^2 - \Psi^2(L)\varepsilon_t^2\right). \tag{63}$$

The AR polynomial, $\tilde{\Psi}(L^k)\prod_{i=1}^{\tilde{r}} \left[1-\zeta_i^k L^k\right]$, is of order $R = k\tilde{r} + \frac{1}{2}kr(r+1)$ in the high frequency time units. The first polynomial of the MA component, $\tilde{\Psi}(L^k)\hat{T}(L)b(L)\Psi^2(L)$, is of order

$$(k-1)\tilde{r} + P + (p+1)(k-1) + q + \frac{1}{2}kr(r-1). \tag{64}$$

And the order of the last MA polynomial (acting on $\varepsilon_{t-i}\varepsilon_{t-j}, i \neq j$) is

$$k\tilde{r} + (p+1)(k-1) + q + \frac{1}{2}kr(r-1). \tag{65}$$

These last two orders are very similar except for $(k-1)\tilde{r}+P$ and $k\tilde{r}$. Notice as well that $\tilde{r} \geq P$ because $\tilde{r} = \max(P, Q)$, implying that $(k-1)\tilde{r}+P \leq k\tilde{r}$. Therefore, the maximum order of the MA polynomial is given by (65). Furthermore, since $r = \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor$, we have $q = kr - (p+1)(k-1)$. Substituting in (65) yields

$$k\tilde{r} + \frac{1}{2}kr(r+1). \tag{66}$$

We have thus shown that the maximum order of the MA polynomial in the low frequency model is $R = \tilde{r} + \frac{1}{2}r(r+1)$. □

9.1 Aggregation of ARMA(1,1)-GARCH(1,1): Order conditions

We perform temporal aggregation of an ARMA(1,1)-GARCH(1,1) model satisfying

$$(1-\phi L)y_t = (1+\theta L)\varepsilon_t, \ \varepsilon_t \sim D(0, h_t) \tag{67}$$

$$h_t = \psi + b h_{t-1} + a\varepsilon_{t-1}^2. \tag{68}$$

We take for simplicity $k = 2$. We focus on the order conditions given in Proposition 6, with the aim of providing an illustration.

The low frequency mean equation is derived substituting the temporal aggregation of (67) into itself, i.e.

$$y_t^* = \phi^2 y_{t-2}^* + \omega_t^*,$$

where $\omega_t^* = \varepsilon_t^* + (\phi+\theta)\varepsilon_{t-1}^* + \phi\theta\varepsilon_{t-2}^* = (1+\phi\theta L^2)\varepsilon_t^* + (\phi+\theta)\varepsilon_{t-1}^*$. We note that $E(\omega_t^*\omega_{t-2j}^*) = 0$, $\forall j > 1$, implying that the low frequency mean equation is also an ARMA(1,1). Regarding the moving average part, we let η be a solution of the system $\frac{\eta}{1+\eta^2} = \frac{E(\omega_t^*\omega_{t-2}^*)}{E((\omega_t^*)^2)}$. Then the aggregate ARMA(1,1) model can be explicitly expressed as

$$y_t^* \quad = \quad \beta y_{t-2}^* + \varepsilon_t^* + \eta\varepsilon_{t-2}^*.$$

The aggregate autoregressive parameter is $\beta = \phi^2$. Note that the low frequency parameters β and η have replaced the high frequency parameters ϕ and θ.

We now have to derive the variance equation, similarly to (68). Let us express the GARCH(1,1) in (68) as an ARMA(1,1) model, setting $\iota_t = \varepsilon_t^2 - h_t$,

$$\varepsilon_t^2 = \psi + (a+b)\varepsilon_{t-1}^2 + \iota_t - b\iota_{t-1}.$$

Substituting the complete expression for ε_{t-1}^2 inside itself,

$$[1-(a+b)^2L^2]\varepsilon_t^2 = (1+a+b)\psi + [1+aL-b(a+b)L^2]\iota_t. \tag{69}$$

We let $\Psi^2(L) = \sum_{i=0}^{1}\psi_i^2 L^i$, and we observe that $(1-(a+b)^2L^2)\Psi^2(L)$ is a polynomial of order 3.

Then, multiplying (69) by $\Psi^2(L)$ and rearranging,

$$\begin{aligned}&[1-(a+b)^2L^2](\varepsilon_T^*)^2 = \Psi^2(1)(1+b+a)\psi + [1+aL-b(a+b)L^2]\Psi^2(L)\iota_t \\ &+[1-(a+b)^2L^2]\left((\Psi(L)\varepsilon_t)^2 - \Psi^2(L)\varepsilon_t^2\right).\end{aligned} \tag{70}$$

The MA structure on the *RHS* is not finite. To overcome this problem, we define $\tilde{\Psi}(L^2) = \sum_{i=0}^{1}\tilde{\psi}_i L^{2i}$. Consequently, $\tilde{\Psi}(L^2)\Psi^2(L)$ is a polynomial of order 3. Again, multiplying (70) by $\tilde{\Psi}(L^2)$,

$$\begin{aligned}&\tilde{\Psi}(L^2)[1-(a+b)^2L^2](\varepsilon_T^*)^2 = \tilde{\Psi}(1)\Psi^2(1)(1+b+a)\psi + \tilde{\Psi}(L^2)[1+aL-b(a+b)L^2]\Psi^2(L)\iota_t \\ &+\tilde{\Psi}(L^2)[1-(a+b)^2L^2]\left((\Psi(L)\varepsilon_t)^2 - \Psi^2(L)\varepsilon_t^2\right).\end{aligned}$$

The AR polynomial on the LHS, $\tilde{\Psi}(L^2)[1-(a+b)^2L^2]$, is of order $\lfloor 4/2 \rfloor = 2$. We can also deduce that the RHS contains only $\iota_t, \ldots, \iota_{t-5}$ and $\varepsilon_{t-i}\varepsilon_{t-j}$ terms ($i \neq j$, i or $j \in (0, \ldots, 5)$). Therefore, the MA order in the temporally aggregated model is equal to $\lfloor 5/2 \rfloor = 2$. We have thus shown that the high frequency ARMA(1,1)-GARCH(1,1) implies a temporally aggregated ARMA(1,1)-GARCH(2,2) model.

9.2 Aggregation of GARCH(1,1) model: Parameters

In this subsection we discuss stock aggregation for a GARCH(1,1) model. We focus on the low frequency implied parameters. For simplicity, we assume as aggregation frequency $k = 2$. The disaggregate GARCH(1,1) model satisfies

$$h_t = \psi + bh_{t-1} + ay_{t-1}^2. \tag{71}$$

Setting $\iota_t = y_t^2 - h_t$, we can express the GARCH as an ARMA model, i.e.

$$y_t^2 = \psi + (a+b)y_{t-1}^2 + \iota_t - b\iota_{t-1}.$$

Substituting the complete expression for y_{t-1}^2 inside the previous equation gives

$$\begin{aligned} & y_t^2 = (1+a+b)\psi + (a+b)^2 y_{t-1}^2 + \iota_t + a\iota_{t-1} - b(a+b)\iota_{t-2} \\ & \Rightarrow (1-(a+b)^2L^2)y_t^2 = (1+a+b)\psi + \iota_t + a\iota_{t-1} - b(a+b)\iota_{t-2}. \end{aligned}$$

We have thus obtained the low frequency ARMA model (in squared observations) corresponding to the original GARCH in (71). If we let $w_t = \iota_t + a\iota_{t-1} - b(a+b)\iota_{t-2}$, we note that $E(w_t w_{t-2j}) = 0$, $\forall j > 1$, implying that the low frequency model is an ARMA(1,1). The low frequency constant equals $(1+a+b)\psi$. The low frequency moving average parameter λ may be recovered solving the system

$$\frac{\lambda}{1+\lambda^2} = \frac{E(w_t w_{t-2})}{E((w_t)^2)} = \frac{b(a+b)}{1+a^2+b^2(a+b)^2},$$

which yields

$$\lambda_{1,2} = \frac{1+2ab^2+b^4+a^2(1+b^2) \pm \sqrt{(1+2ab^3+b^4+a^2(1+b^2))^2 - 4b^2(a+b)^2}}{2b(a+b)}. \tag{72}$$

To calculate the low frequency autoregressive parameter, we note that the low frequency model may be expressed in disaggregate time units as

$$y_t^2 = (1+a+b)\psi + (a+b)^2 y_{t-2}^2 + \iota_t + \lambda\iota_{t-2}.$$

At the same time, definition (58) holds. Hence, the best linear predictor is

$$P[y_t^2 | y_{t-2}, y_{t-4}, \ldots] = (1+a+b)\psi + ((a+b)^2 + \lambda)y_{t-2}^2, \tag{73}$$

reminding that ι_{t-2} is a function of $y_{t-2}, y_{t-4}, \ldots$. Therefore, the low frequency autoregressive parameter equals $(a+b)^2 + \lambda$, with λ satisfying (72).

10 EMPIRICAL APPLICATION: BELGIAN PUBLIC DEFICIT

In this section we present an empirical application of the techniques so far surveyed. We apply the temporal aggregation methodology to the Belgian public deficit series. We use the *net balance to be* *financed*- federal deficit in sort. This is the definition used by the Belgian Federal Public Service Finance, www.minfin.fgov.be. This concept is very different from the ESA95 (or excessive deficit procedure) definitions of net lending or net borrowing. It is on cash basis, contrary to be ESA95 deficit that is on accrual basis. It means that it is influenced by financial transactions (e.g. privatization receipts) that have no effect on the ESA95 deficit figures. Another crucial difference is that the federal deficit is released monthly while the ESA95 is released annually as it is the result of the annual national account that follows ESA95.

The federal deficit is used by the federal Treasury for instance in order to monitor debt management. It limited to (federal government) Treasury operations; as a consequence, it does not take into account operations of other federal institutions, social security institutions, regions and communities or local authorities. Data range from January 1981 until December 2001, in real terms.[14] The data set used contains 252 monthly observations. Our source is the National Bank of Belgium.

Figure 3 shows the Belgian cash deficit series, in real terms, at different frequencies. Top left panel displays the series at the monthly frequency, top right panel at the quarterly frequency, while the bottom panel presents the annual series. It can be clearly seen that monthly and quarterly observations have a very strong seasonal pattern. The seasonality is due to the intra-annual instalments corresponding to the tax collection and payments. For instance, advance payments by companies are received in the months April, July, October and December. As a result, in these months the federal deficit is generally positive for the federal government. As these instalments are intra-annual, the annual series does not reflect any seasonal pattern. On the basis of the explanations given in Section 7, the annual seasonality vanishes because $s = 12$ and $k = 12$. Related with, notice that it is not clear whether the monthly series, and even the quarterly series, have a unit root in levels or not. But it is very evident that the annual series has a unit root. As we shall see, the monthly series has a seasonal unit root. This dovetails perfectly with the theory. If $k = s = 12$, the seasonal unit root at the monthly frequency becomes a regular unit root at the annual frequency.

[FIGURE 3 ABOUT HERE]

We perform several estimation exercises. We start estimating a monthly model.[15] We then aggregate it quarterly and annually.[16] To make a comparison, we estimate directly the same quarterly and annual models from the quarterly and annual data sets.

The monthly series can be represented by an ARIMA$(0,0,1) \times (0,1,1)_{12}$ model with mean:

$$(1 - L^{12})y_t = \hat{\mu} + (1 + \hat{\theta}L)(1 + \hat{\Theta}L^{12})\varepsilon_t, \tag{74}$$

that is a model with a unit root in seasonality. The estimated parameters are shown in Table 2.

[TABLE 2 ABOUT HERE]

[14]The original cash deficit time series, in nominal terms, has been deflated dividing by the monthly Belgian CPI, base year 1996.

[15]We use the package TRAMO to fit the best model. TRAMO (Time Series Regression with ARIMA Noise, Missing Observations, and Outliers) performs estimation, forecasting, and interpolation of regression models with missing observations and ARIMA errors, in the presence of several types of outliers. TRAMO estimates a battery of models, including the outliers analysis, and selects the best model using the Bayes Information Criteria. For a brief description of the program, consult the webpage http://www.bde.es/servicio/software/tswe.htm.

[16]The Matlab codes for temporal aggregation of several ARIMA models with $k = 12$ are available on the homepage of David Veredas at http://www.ecares.org.

First, we aggregate the monthly ARIMA$(0,0,1)\times(0,1,1)_{12}$ model in (74) to quarterly periodicity. We multiply both sides of (74) by $\sum_{i=0}^{2} L^j$, obtaining:

$$\left(1-L^{12}\right)\sum_{j=0}^{2} y_{t-j} = 3\hat{\mu} + \left(1+\hat{\theta}L\right)\left(1+\hat{\Theta}L^{12}\right)\sum_{j=0}^{2}\varepsilon_{t-j}. \tag{75}$$

Letting $B = L^3$ to operate on the aggregate time unit T, this quarterly aggregated model may be expressed as an ARIMA$(0,0,1)\times(0,1,1)_4$

$$\left(1-B^4\right) y_T^* = M + (1+\eta_1 B)\left(1+E_1B^4\right)\varepsilon_T^*, \tag{76}$$

where the aggregate parameters E_1 and $\sigma^2_{\varepsilon^*}$ are determined solving the system:

$$\begin{aligned}
(1+\eta_1^2+E_1^2+\eta_1^2E_1^2)\sigma^2_{\varepsilon^*} &= (1+2(1+\hat{\theta})^2+\hat{\theta}^2+\hat{\Theta}^2+2\hat{\Theta}^2(1+\hat{\theta})^2+\hat{\theta}^2\hat{\Theta}^2)\hat{\sigma}^2_{\varepsilon}\\
(1+E_1^2)\eta_1\sigma^2_{\varepsilon^*} &= \hat{\theta}(1+\hat{\Theta}^2)\hat{\sigma}^2_{\varepsilon}\\
\eta_1E_1\sigma^2_{\varepsilon^*} &= \hat{\theta}\hat{\Theta}\hat{\sigma}^2_{\varepsilon}\\
(1+\eta_1^2)E_1\sigma^2_{\varepsilon^*} &= \hat{\Theta}(1+2(1+\hat{\theta})^2+\hat{\theta}^2)\hat{\sigma}^2_{\varepsilon}.
\end{aligned}$$

Notice that everything on the RHS here above is known as it has already been estimated. The equation

$$\frac{(1+E_1^2)\eta_1\sigma^2_{\varepsilon^*}}{\eta_1E_1\sigma^2_{\varepsilon^*}} = \frac{\hat{\theta}(1+\hat{\Theta}^2)\hat{\sigma}^2_{\varepsilon}}{\hat{\theta}\hat{\Theta}\hat{\sigma}^2_{\varepsilon}} \Rightarrow \frac{(1+E_1^2)}{E_1} = \frac{(1+\hat{\Theta}^2)}{\hat{\Theta}}$$

immediately delivers $E_1=\hat{\Theta}$. Once E_1 is recovered, the system above simplifies in:

$$\begin{aligned}
(1+\eta_1^2)\sigma^2_{\varepsilon^*} &= (1+2(1+\hat{\theta})^2+\hat{\theta}^2)\hat{\sigma}^2_{\varepsilon}\\
\eta_1\sigma^2_{\varepsilon^*} &= \hat{\theta}\hat{\sigma}^2_{\varepsilon}.
\end{aligned}$$

And the aggregate inferred parameters are reported in Table 2.

Second, we perform temporal aggregation into annual frequency. To do this, as already extensively detailed, we multiply both sides of (74) by $\sum_{i=0}^{11} L^j$. In this way the model in (74) is expressed in terms of the aggregate annual periodicity,

$$\left(1-L^{12}\right)\sum_{j=0}^{11} y_{t-j} = 12\hat{\mu} + (1+\hat{\theta}L)\left(1+\hat{\Theta}L^{12}\right)\sum_{j=0}^{11}\varepsilon_{t-j}. \tag{77}$$

If we let $B=L^{12}$ operate on the annual temporal index $T=1,2,\ldots$, the aggregate model in (77) corresponds to an ARIMA(0,1,2):

$$(1-B)\, y_T^* = M + \left(1+\eta_1B+\eta_2B^2\right)\varepsilon_T^*. \tag{78}$$

Note that in the annual model the seasonal unit root becomes a normal unit root. The aggregate parameters η_1, η_2 and $\sigma^2_{\varepsilon^*}$ are determined by the system:

$$\begin{aligned}
\left(1+\eta_1^2+\eta_2^2\right)\sigma^2_{\varepsilon^*} &= [12+12\hat{\theta}^2+12\hat{\Theta}^2+12\hat{\theta}^2\hat{\Theta}^2+2\hat{\theta}(11+\hat{\Theta}+11\hat{\Theta}^2)]\hat{\sigma}^2_{\varepsilon}\\
(\eta_1+\eta_1\eta_2)\,\sigma^2_{\varepsilon^*} &= \left[\left(\hat{\theta}+\hat{\Theta}\right)+11\left(1+\hat{\theta}\right)^2\hat{\Theta}+\hat{\theta}\hat{\Theta}\left(\hat{\theta}+\hat{\Theta}\right)\right]\hat{\sigma}^2_{\varepsilon}\\
\eta_2\sigma^2_{\varepsilon^*} &= \hat{\theta}\hat{\Theta}\hat{\sigma}^2_{\varepsilon}.
\end{aligned}$$

The aggregate inferred parameters are in Table 2.

We also estimate the quarterly and annual models directly, from the aggregate observations (quarterly and annual data sets), that are shown in Table 3.

[TABLE 3 ABOUT HERE]

Some conclusions may be drawn. We divide them in *within*, meaning the analysis of the monthly and temporally aggregated models (i.e. within Table 2), and *between*, meaning the differences between the models in Tables 2 and 3.

Within conclusions. First, the constants perfectly reflect the spirit of the aggregation technique: the quarterly and annual constants are exactly 3 and 12 times the monthly constant. Second, the annual model does not have seasonal MA components and it becomes an MA(2). Moreover, even if we aggregate 12 periods, the inferred parameters are still large, meaning that there is significant autocorrelation in the annual series. Last, the inferred residual variance increases over time. This makes sense, since cash deficit is a flow variable and hence the more we aggregate, the larger the deficit becomes (as this is reflected in Figure 3, where the monthly deficit roughly ranges from -0.04 to 0.04 while the annual deficit ranges from -0.20 to 0).

Between conclusions. First, the constants are very similar, meaning that there is a coherence and that the temporally aggregated models are doing their job. Second, by contrast, the MA parameters differ, especially for the annual frequency. A possible cause is that the number of annual observations is 21, comparing with the 252 of the aggregated model, suggesting that asymptotic properties merely work. Moreover, recall that the transfer function from the monthly model to the quarterly and annual aggregated models is deterministic. This is a fundamental feature. It means that all the efficiency (due to 252 observations, 12 times more than 21) of the monthly estimates is transferred to the annual ones, with no loss whatsoever. Regarding the residual variance, finally, direct estimation and temporal aggregation produce comparable results.

11 CONCLUDING REMARKS

The primary purpose of this paper is to provide a comprehensive analysis of temporal aggregation, a topic of fundamental importance in econometrics and in time series analysis. The main question we try to answer can be expressed as follows: if, it is known that an observed stochastic process may be described by a linear time series model, what is the appropriate model for the same process observed at different frequencies? To answer this question, we address a number of issues that arise whenever the sample frequency does not correspond to the 'natural' frequency (i.e. the frequency at which the process is supposed to be generated). In general, the techniques discussed are very flexible and can therefore be used in a variety of different fields.

We review the econometric methodology for univariate time series temporal aggregation. We start with simple AR models, ending up with more difficult ARIMA models with seasonality, independent exogenous variables and GARCH effects. In Table 4, to summarize, we provide a general overview of the results obtained for the class of linear processes presented throughout the paper.

[TABLE 4 ABOUT HERE]

As it is evident from Table 4, the most striking departures with respect to the original model specifications are linked to the MA orders. Numerical procedures are very often required to determine the MA and AR aggregate parameters, as we show in several example-cases discussed in this article.

Lines for further developments stretch both in theoretical both in empirical directions. On a theoretical ground, issues like multivariate models, nonlinearities, long memory, random aggregation, time continuous or state space representations are of great relevance. All these issues show interesting features under temporal aggregation that will be undoubtedly worth analyzing in the future. On a more applied framework, the consequences of temporal aggregation on parameter estimation and forecasting have not been investigated. Exemplifying, if disaggregate data are available and we are interested in estimating parameters, the maximum likelihood estimators may be based on the disaggregate model or on the aggregate model. How carefully measuring the information loss in estimation due to aggregation? If, similarly, we are interested in forecasting aggregate observations, the forecast function can be based either on the disaggregate series either on the aggregate series. Does it exist a loss in efficiency through aggregation? Is it possible to precisely quantify it? These are still partially open questions left to future research.

References

[1] Abraham B. 1982. Temporal aggregation and time series, *International Statistical Review*, **50**: 285-291.

[2] Amemiya T., Wu R.Y. 1972. The effect of aggregation on prediction in the autoregressive model, *Journal of the American Statistical Association*, **67**: 628-632.

[3] Aranzana M., Veredas D. 2005. Missing the unit roots. Or what happens if you miss the missing values. Mimeo.

[4] Bollerslev T. 1986. Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, **31**: 307-327.

[5] Box G.E.P., Jenkins G.M., Reinsel G.C. 1994. Time series analysis: forecasting and control, 3rd edition, Prentice Hall (eds.). New York.

[6] Breitung J., Swanson N.R. 2002. Temporal aggregation and spurious instantaneous causality in multiple time series models, *Journal of Time Series Analysis*, **23**: 651-665.

[7] Drost F.C., Werker B.J.M. 1996. Closing the GARCH gap: continuous GARCH modelling, *Journal of Econmetrics*, **74**: 31-57.

[8] Drost F.C., Nijman T.E. 1993. Temporal aggregation of GARCH processes, *Econometrica*, **61**: 909-927.

[9] Drost F.C. 1994. Temporal aggregation of time series. In *Econometric Analysis of Financial Markets*, Kaehler and Kugler (eds.), Physica-Verlag: 11-21.

[10] Engle R.F. 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of the U.K. inflation, *Econometrica*, **50**: 987-1008.

[11] Granger C.W.J. 1980. Long memory relationships and the aggregation of dynamic models, *Journal of Econometrics*, **14**: 227-280.

[12] Granger C.W.J. 1987. Implications of aggregation with common factors, *Econometric Theory*, **3**: 208-222.

[13] Granger C.W.J. 1990. Aggregation of time series variables: a survey. *J.R. Statistical Society A*, **139**: 246-2257.

[14] Granger C.W.J, Lee T-H. 1993. The effect of aggregation on nonlinearity. In *Advances in Statistical Analysis and Statistical Computing*, Mariano R. (eds.), Vol. 3, JAI Press Inc.

[15] Granger C.W.J, Morris M.J. 1976. Time Series Modelling and Interpretation. In *Advances in Statistical Analysis and Statistical Computing*, Mariano R. (eds.), Vol. 3, JAI Press Inc.

[16] Hafner C.M., Rombouts J.V.K. 2003. Estimation of temporally aggregated multivariate GARCH models. CORE Discussion Paper 2003/73, Université catholique de Louvain, Belgium.

[17] Hafner C.M. 2004. Temporal aggregation of multivariate GARCH processes. Econometric Institute Report 2004-29, Erasmus University Rotterdam, The Netherlands.

[18] Harvey A.C. 1981. Time series models. Philip Allan (eds.). Oxford.

[19] Hotta L.K., Cardoso Neto J. 1993. The effect of aggregation on prediction in autoregressive integrated moving average models, *Journal of Time Series Analysis*, **14**: 261-269.

[20] Hotta L.K., Valls Pereira P.L., Otta R. 2004. Effect of outliers on forecasting temporally aggregated flow variables, *Sociedad de Estadística e Investigación Operativa Test*, **13**: 371-402.

[21] Jorda O., Marcellino M. 2004. Time-scale transformations of discrete time series processes, *Journal of Time Series Analysis*, **25**: 873-894.

[22] Koreisha S.G., Fang Y. 2004. Updating ARMA predictions for temporal aggregates, *Journal of Forecasting*, **23**: 275-296.

[23] Lütkepohl H. 1984. Forecasting contemporaneously aggregated vector ARMA processes, *Journal of Business and Economic Statistics*, **2**: 201-214.

[24] Lütkepohl H. 1986. Comparison of predictors for temporally and contemporaneously aggregated time series, *International Journal of Forecasting*, **2**: 461-475.

[25] Lütkepohl H. 1987. Forecasting aggregated vector ARMA processes. In *Lecture Notes in Economics and Mathematical Systems*, Springer Verlag (eds.). Berlin, New York.

[26] Marcellino M. 1999. Some consequences of temporal aggregation in empirical analysis, *Journal of Business and Economic Statistics*, **17**: 129-136.

[27] Moulin L., Salto M., Silvestrini A., Veredas D. 2004. Using intra annual information to forecast the annual State deficits. The case of France. CORE Discussion Paper 2004/48, Université catholique de Louvain, Belgium.

[28] Nelson, D.B. 1990. ARCH models as diffusion approximations, *Journal of Econometrics*, **45**: 7-38.

[29] Nijman T.E., Palm F.C. 1990. Predictive accuracy gain from disaggregate sampling in ARIMA models, *Journal of Business and Economic Statistics*, **8**: 405-415.

[30] Nijman T.E., Palm F.C. 1990. Parameter identification in ARMA processes in the presence of regular but incomplete sampling, *Journal of Time Series Analysis*, **11**: 239-248.

[31] Palm F.C., Nijman T.E. 1984. Missing observations in the dynamic regression model, *Econometrica*, **52**: 1415-1435.

[32] Proietti T. 2005. On the estimation of nonlinearly aggregated mixed models, accepted for publication in *Journal of Computational and Graphical Statistics*.

[33] Stram D.O., Wei W.W.S. 1986. Temporal aggregation in the ARIMA process, *Journal of Time Series Analysis*, **7**: 279-292.

[34] Tiao G.C. 1972. Asymptotic behaviour of temporal aggregates of time series, *Biometrika*, **59**: 525-531.

[35] Wei W.W.S. 1978. Some consequences of temporal aggregation in seasonal time series models. In *Seasonal Analysis of Economic Time Series*, Zellner A. (eds.). US Department of Commerce, Bureau of the Census: Washington, DC.

[36] Weiss A. 1984. Systematic sampling and temporal aggregation in time series models, *Journal of Econometrics*, **26**: 271-281.

[37] Weiss A. 1986. Asymptotic theory for ARCH models: estimation and testing, *Econometric Theory*, **2**: 107-131.

A APPENDIX

A.1 Stock aggregation of ARMA models: Order conditions

Let us hereafter focus on stock aggregation. The underlying process follows the ARMA(p,q) model in (13): $\phi(L)y_t = \theta(L)\varepsilon_t$, where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ and $\theta(L) = 1 + \theta_1 L + \ldots + \theta_q L^q$ are the AR and MA polynomials of length p and q, respectively. The variable ε_t is an independent white noise error term with mean to equal zero and constant variance σ_ε^2. Given the model in (13), suppose that only the kth, $2k$th, $3k$th, ... terms are observed. This corresponds to a stock aggregation scheme, where k is the integer interval at which we sample the low frequency data. The aim is to find an appropriate model for the aggregate stock variable $y_{T=tk}$. Consider the proposition that follows.

Proposition 7 *The sampled series of y_t as specified in model (13), denoted $\{y_T,\ T \in \mathbb{Z}\}$ (where $y_T = y_{kt}$ and k is the integer interval at which we observe the low frequency data), may be represented by an ARMA(p,r) process where r, the maximum order of the sampled moving average polynomial, is*

$$r = \left\lfloor \frac{p(k-1)+q}{k} \right\rfloor . \tag{79}$$

Proof. To prove this, consider a generic polynomial $S(L)$ of length h whose coefficients are $1, s_1, \ldots, s_h$. Multiplying each side of (13) by $S(L)$, we obtain:

$$S(L)\phi(L)y_t = S(L)\theta(L)\varepsilon_t.$$

On the LHS of (80), the product $S(L)\phi(L)$ defines the autoregressive polynomial for the aggregate series. This autoregressive polynomial has to sample y_t in (13) every k terms, producing a sequence $\{y_T = y_{kt},\ T \in \mathbb{Z}\}$. Brewer (1973) expresses this requirement through the following equation in matrix form, i.e.

$$S(L)\phi(L) = \begin{pmatrix} 1 & 0 & \ldots & 0 \\ s_1 & 1 & \ldots & 0 \\ s_2 & s_1 & 1 & 0 \\ \ldots & \ldots & \ldots & \ldots \\ s_p & s_{p-1} & \ldots & 1 \\ s_h & s_{h-1} & \ldots & s_{h-p} \\ 0 & s_h & \ldots & \ldots \\ 0 & 0 & s_h & \ldots \\ 0 & 0 & 0 & s_h \end{pmatrix} \begin{pmatrix} 1 \\ -\phi_1 \\ \ldots \\ -\phi_p \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \ldots \\ \beta_1 \\ 0 \\ 0 \\ \ldots \\ \beta_2 \\ 0 \\ 0 \\ \ldots \\ \beta_s \end{pmatrix}. \tag{80}$$

In (80), the first matrix on the LHS, $S(L)$, has size $(p+h+1) \times (p+1)$. The second one on the LHS, corresponding to the AR polynomial $\phi(L)$, has size $(p+1) \times 1$. Hence, the product $S(L)\phi(L)$ has dimension $(p+h+1) \times 1$. The RHS matrix, instead, has size $(rk+1) \times 1$. No zeros are needed below β_s, otherwise the last equation would be $-s_h\phi_p = 0$, which is a contradiction since both s_h both ϕ_p are non-zero by definition.

It is now necessary to obtain the unknown orders h and r (AR polynomial length in the aggregate series) as a function of k and p (AR polynomial length in the original series). The first matrix on the LHS is $(p+h+1) \times 1$, by construction. The RHS matrix is $(rk+1) \times 1$. Hence, following Brewer, the first restriction is

$$p + h + 1 = rk + 1 \Rightarrow p + h = rk. \tag{81}$$

In addition, there are $s_1, s_2, \ldots, s_h$ unknown coefficients in the $S(L)$ polynomial to match with $r(k-1)$ conditions that have been imposed. On the RHS matrix, indeed, there are exactly $r(k-1)$ zeros corresponding to the null coefficients of the RHS matrix, by definition. Therefore, according to Brewer, the second restriction is

$$h = r(k-1) \Rightarrow rk - p = rk - r \Rightarrow p = r. \tag{82}$$

Then, the AR polynomial length in the underlying model expressed by equation (13) equals exactly the AR polynomial length in the aggregate series. The autoregressive order of the aggregate model is unchanged at p. Substituting (81) in (82), we get $h = p(k-1)$. Hence, the unknown length of $S(L)$ polynomial corresponds to $p(k-1)$.

Let us now consider the MA polynomial on the RHS of (80). We already know that the moving average polynomial $\theta(L)$ has length q. As a consequence, the length of $S(L)\theta(L)$ is

$$h + q = r(k-1) + q = p(k-1) + q. \tag{83}$$

This has a clear impact on the autocovariance structure of the aggregate series

$$E[\eta(L)\varepsilon_t \eta(L)\varepsilon_{t-kr}],$$

with $k \in N_0$. For $kr \leq p(k-1) + q$, we have:

$$E[\eta(L)\varepsilon_t \eta(L)\varepsilon_{t-kr}] \neq 0.$$

Else, for $kr \geq p(k-1) + q$:

$$E[\eta(L)\varepsilon_t \eta(L)\varepsilon_{t-kr}] = 0. \tag{84}$$

This means that, in equation (80), the whole RHS may be replaced by a MA aggregate polynomial of length $\lfloor k^{-1}(p(k-1)+q) \rfloor$. Consequently, the underlying model in equation (13), sampled every kth terms, may be appropriately represented by an $ARMA(p, \lfloor k^{-1}(p(k-1)+q) \rfloor)$ model, as reported in (79). □

A.2 Stock aggregation of ARIMA models: Order conditions

We deal with stock variables in model (25). Hence, we face a stock aggregation scheme in the presence of ARIMA models.

Proposition 8 *The sampled series of y_t as specified in model (25), denoted $\{y_T,\ T \in \mathbb{Z}\}$ (where $y_T = y_{kt}$ and k is the integer interval at which we observe the low frequency data), may be represented by an ARIMA(p,d,r) process where r, the order of the sampled moving average polynomial, is*

$$r = \left\lfloor \frac{p(k-1) + d(k-1) + q}{k} \right\rfloor. \tag{85}$$

Proof. We give a complete proof of the order condition in (85). As suggested by Weiss (1984), we apply to both sides of the ARIMA(p,d,q) model in (25) the operator $\bar{S}(L)$ listed in (35), obtaining

$$\prod_{j=1}^{p} \left[1 - \delta_j^k L^k\right] (1 - L^k)^d y_t = \prod_{j=1}^{p} \left[\frac{1 - \delta_j^k L^k}{1 - \delta_j L}\right] \left[\frac{1 - L^k}{1 - L}\right]^d \left(\sum_{j=0}^{q} \theta_j L^j\right) \varepsilon_t.$$

Therefore, filtering with the operator in (35) an ARIMA(p,d,q) model with stock variables, we have shown that the sampled series of y_t (denoted $\{y_T,\ T \in \mathbb{Z}\}$) may be represented, letting $B = L^k$, by an ARIMA(p,d,r)

$$\prod_{j=1}^{p} \left[1 - \delta_j^k B\right] (1 - B)^d y_T = (1 + \eta_1 B + ... + \eta_r B^r)\varepsilon_T, \tag{86}$$

with

$$r = k^{-1} \lfloor p(k-1) + d(k-1) + q \rfloor.$$

In sum, the initial ARIMA(p,d,q) aggregates to an ARIMA(p,d,r), where $r = [(p + d) + (q - p - d)/k]$. This latter value represents the maximum length of the MA polynomial in the aggregate model. We remark that, also in this case, the AR order is unchanged by the sampling scheme. In addition, the roots of the AR polynomial in the sampled series are the kth powers of the inverted roots of the AR polynomial in the disaggregate one. Thus, for large values of the sampling interval k, the AR aggregate coefficients decrease in size. □

A.3 Stock aggregation and seasonality: Order conditions

Dealing with stock variables in (29), we report the following proposition for a stock aggregation scheme in the presence of ARIMA models with seasonality.

Proposition 9 *The sampled series of y_t as specified in model (29), denoted $\{y_T,\ T \in \mathbb{Z}\}$ (where $y_T = y_{kt}$ and k is the integer interval at which we observe the low frequency data), may be represented by an $ARIMA(p,d,r) \times (P,D,R)_{s^*}$ where r, the maximum order of the sampled regular moving average polynomial is*

$$r = \left\lfloor \frac{p(k-1) + d(k-1) + q}{k} \right\rfloor, \tag{87}$$

and R, the maximum order of the sampled moving average in seasonality is

$$R = \left\lfloor \frac{(P+D)s^* k + (Q-P-D)s}{k} \right\rfloor. \tag{88}$$

Proof. We now prove the order conditions reported in (87) and (88). Let $\tau_1^s, \ldots, \tau_P^s$ be the reciprocals of the roots of $\Phi(L^s)$ polynomial, such that $\Phi(L^s) = \prod_{i=1}^{P} [1 - (\tau_i L)^s]$. In the case of stock variables, let us consider again the operators $\bar{S}(L)$ in (35) and $A(L)$ in (30). Applying the product of these two operators to both sides of (29), we get

$$\bar{S}(L)A(L)\phi(L)\Phi(L^s)(1-L)^d(1-L^s)^D y_t = \bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t, \tag{89}$$

which can be developed as follows

$$\begin{aligned} &\prod_{j=1}^{p}\left[\frac{1-\delta_j^k L^k}{1-\delta_j L}\right]\left[\frac{1-L^k}{1-L}\right]^d \prod_{i=1}^{P}\left[\frac{1-(\tau_i L)^{ks^*}}{1-(\tau_i L)^s}\right]\left[\frac{1-L^{ks^*}}{1-L^s}\right]^D \\ &\times \phi(L)\Phi(L^s)(1-L)^d(1-L^s)^D y_t = \bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t. \end{aligned} \tag{90}$$

The complete expressions for $\phi(L)$ and $\Phi(L^s)$ are $\Phi(L^s) = \prod_{i=1}^{P} [1 - (\tau_i L)^s]$ and $\phi(L) = \prod_{j=1}^{p} [1 - \delta_j L]$. Substituting them in (90) we get

$$\begin{aligned} &\prod_{j=1}^{p}\left[\frac{1-\delta_j^k L^k}{1-\delta_j L}\right]\left[\frac{1-L^k}{1-L}\right]^d \prod_{i=1}^{P}\left[\frac{1-(\tau_i L)^{ks^*}}{1-(\tau_i L)^s}\right]\left[\frac{1-L^{ks^*}}{1-L^s}\right]^D \\ &\times \prod_{j=1}^{p}[1-\delta_j L]\prod_{i=1}^{P}[1-(\tau_i L)^s](1-L)^d(1-L^s)^D y_t \\ &= \bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t. \end{aligned}$$

Simplifying, we obtain

$$\begin{aligned} &\prod_{j=1}^{p}\left[1-\delta_j^k L^k\right]\left[\frac{1-L^k}{1-L}\right]^d \prod_{i=1}^{P}\left[1-(\tau_i L)^{ks^*}\right]\left[\frac{1-L^{ks^*}}{1-L^s}\right]^D \\ &\times (1-L)^d(1-L^s)^D y_t = \bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t. \end{aligned}$$

Manipulating further this expression, we have

$$\begin{aligned} &\prod_{j=1}^{p}\left[1-\delta_j^k L^k\right]\prod_{i=1}^{P}\left[1-(\tau_i L)^{ks^*}\right](1-L^k)^d(1-L^{ks^*})^D y_t \\ &= \bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t. \end{aligned}$$

To check the order of the MA polynomial is exactly the one in (87), we must develop the RHS term as

$$\prod_{j=1}^{p}\left[\frac{1-\delta_j^k L^k}{1-\delta_j L}\right]\left[\frac{1-L^k}{1-L}\right]^d\prod_{i=1}^{P}\left[\frac{1-(\tau_i L)^{ks^*}}{1-(\tau_i L)^s}\right]\left[\frac{1-L^{ks^*}}{1-L^s}\right]^D\theta(L)\Theta(L^s)\varepsilon_t$$
$$=\prod_{j=1}^{p}\left[\frac{1-\delta_j^k L^k}{1-\delta_j L}\right]\left[\sum_{j=0}^{d(k-1)}L^j\right]\prod_{i=1}^{P}\left[\frac{1-(\tau_i L)^{ks^*}}{1-(\tau_i L)^s}\right]\left[\frac{1-L^{ks^*}}{1-L^s}\right]^D\theta(L)\Theta(L^s)\varepsilon_t. \tag{91}$$

As it is possible to deduce from (91), the order of the regular moving average polynomial is $\lfloor d(k-1)+p(k-1)+q\rfloor$, while the order of the seasonal moving average polynomial is $\lfloor (P+D)s^*k+(Q-P-D)s\rfloor$.

Thus, the sampled series $\{y_T,\ T\in\mathbb{Z}\}$ follows an $ARIMA(p,d,r)\times(P,D,R)_{s^*}$ process, that can be represented letting $B=L^k$ as

$$\prod_{j=1}^{p}\left[1-\delta_j^k B\right]\prod_{i=1}^{P}\left[1-\tau_i^k B^{s^*}\right](1-B)^d(1-B^{s^*})^D y_T$$
$$=(1+\eta_1 B+\ldots+\eta_r B^r)(1+E_1 B+\ldots+E_R B^R)\varepsilon_T, \tag{92}$$

where r and R are the orders reported in (87) and in (88). Picking s^* as previously indicated ($s^*=s/k$), this implies that the sampled series follows a seasonal $ARIMA(p,d,r)\times(P,D,0)_{s^*}$ model. □

A.4 Stock aggregation of GARCH models: Order conditions

In what follows we discuss stock aggregation of the ARMA(p,q)-GARCH(P,Q) model in (56).

Proposition 10 *The sampled series of y_t as specified in (56), denoted $\{y_T,\ T\in\mathbb{Z}\}$ (where $y_T=y_{kt}$ and k is the integer interval at which we observe the low frequency data), may be represented by an ARMA(p,r) with weak GARCH(R,R) errors where*

$$r = \left\lfloor\frac{p(k-1)+q}{k}\right\rfloor$$
$$R = \tilde{r}+\frac{1}{2}r(r+1),$$

and where $\tilde{r}=\max(P,Q)$.

Proof. As a first step, we start with the conditional mean, applying to both sides of $\phi(L)y_t=\theta(L)\varepsilon_t$ the operator

$$S(L)=\prod_{i=1}^{p}\left[\frac{1-\delta_i^k L^k}{1-\delta_i L}\right], \tag{93}$$

where δ_i, $i=1,\ldots,p$ represent the inverted roots of $\phi(L)$ polynomial, assumed to be stationary. We obtain as a result an ARMA(p,r) model

$$\prod_{i=1}^{p}(1-\delta_i^k L^k)y_t=S(L)\theta(L)\varepsilon_t,$$

which may be expressed in aggregate time units as

$$\beta(B)y_T=\eta(B)\varepsilon_T,$$

with $r=\left\lfloor\frac{p(k-1)+q}{k}\right\rfloor$.

The second step of the proof is almost identical to the flow aggregation case. We express the GARCH model (57) as an ARMA, as done in (59).

The covariance structures of $\eta(B)\varepsilon_T$ and $S(L)\theta(L)\varepsilon_t$ are the same. Therefore $\eta(L^k)\varepsilon_T = S(L)\theta(L)\varepsilon_t$. Or $\varepsilon_T = \eta(L^k)^{-1}S(L)\theta(L)\varepsilon_t = \Psi(L)\varepsilon_t$, where $\Psi(L) = \eta(L^k)^{-1}S(L)\theta(L)$ is a polynomial of infinite order. We define $\hat{T}(L) = \prod_{i=1}^{\tilde{r}} \left[\frac{1-\zeta_i^k L^k}{1-\zeta_i L}\right]$ and $\Psi^2(L) = \sum_{i=0}^{k-1} \psi_i^2 L^i$ as for flow aggregation. We multiply (59) by $\hat{T}(L)$ times $\Psi^2(L)$. This product of polynomials links the GARCH models for ε_t and ε_T, exactly as shown in (60).

Equation (60) can be rewritten in terms of $(\varepsilon_T)^2$, in the same fashion of (61). And in (61) we have a GARCH structure for the low frequency model. However, the MA part is not finite. We multiply (61) by the finite polynomial $\eta(L^k)\Psi(L) = \eta(L^k)\eta(L^k)^{-1}S(L)\theta(L) = S(L)\theta(L)$, of order $p(k-1)+q$. Therefore $\sum_{j=0}^{r} \eta_j \psi_{i-jk} = 0,\ for\ i > p(k-1)+q$. This difference equation determines $\sum_{j=0}^{\frac{1}{2}r(r+1)} \tilde{\psi}_j \psi_{i-jk}^2 = 0,\ for\ i > p(k-1)+q+\frac{1}{2}kr(r-1)$, a difference equation in ψ^2 of order $\frac{1}{2}r(r+1)$. Defining $\tilde{\Psi}(L^k) = \sum_{i=0}^{\frac{1}{2}kr(r+1)} \tilde{\psi}_i L^{ik}$, we can finally observe that $\tilde{\Psi}(L^k)\Psi^2(L)$ is a polynomial of order $p(k-1)+q+\frac{1}{2}kr(r-1)$.

Multiplying (61) by $\tilde{\Psi}(L^k)$ yields

$$\tilde{\Psi}(L^k)\prod_{i=1}^{\tilde{r}}\left[1-\zeta_i^k L^k\right](\varepsilon_T)^2 = \tilde{\Psi}(1)\hat{T}(1)\Psi^2(1)\psi +$$

$$\tilde{\Psi}(L^k)\hat{T}(L)b(L)\Psi^2(L)\iota_t + \tilde{\Psi}(L^k)\prod_{i=1}^{\tilde{r}}\left[1-\zeta_i^k L^k\right]\left((\Psi(L)\varepsilon_t)^2 - \Psi^2(L)\varepsilon_t^2\right). \tag{94}$$

The AR polynomial, $\tilde{\Psi}(L^k)\prod_{i=1}^{\tilde{r}}\left[1-\zeta_i^k L^k\right]$, is of order $R = k\tilde{r} + \frac{1}{2}kr(r+1)$ in the high frequency time units. The first polynomial of the MA component, $\tilde{\Psi}(L^k)\hat{T}(L)b(L)\Psi^2(L)$, is of order $(k-1)\tilde{r}+P+p(k-1)+q+\frac{1}{2}kr(r-1)$. And the order of the last MA polynomial (acting on $\varepsilon_{t-i}\varepsilon_{t-j}, i \neq j$) is

$$k\tilde{r} + p(k-1) + q + \frac{1}{2}kr(r-1). \tag{95}$$

These last two orders are very similar except for $(k-1)\tilde{r}+P$ and $k\tilde{r}$. Moreover, as we know, $\tilde{r} \geq P$, since $\tilde{r} = \max(P,Q)$. This implies that $(k-1)\tilde{r}+P \leq k\tilde{r}$. Therefore, the maximum order of the MA polynomial is given by (95). In addition, $q = kr - (p+1)(k-1)$. Substituting in (95) gives

$$k\tilde{r} + \frac{1}{2}kr(r+1). \tag{96}$$

We have thus shown that the maximum MA order in the low frequency model is $R = \tilde{r} + \frac{1}{2}r(r+1)$. □

Table 1: Temporal aggregation effect in seasonality

k	$\frac{k}{s}$	s^*	Cycle length
2	$\frac{2}{12} = \frac{1}{6}$	6	1 year
3	$\frac{3}{12} = \frac{1}{4}$	4	1 year
4	$\frac{4}{12} = \frac{1}{3}$	3	1 year
5	$\frac{5}{12}$	12	5 years
6	$\frac{6}{12} = \frac{1}{2}$	2	1 year
7	$\frac{7}{12}$	12	7 years
8	$\frac{8}{12} = \frac{2}{3}$	3	2 years
9	$\frac{9}{12} = \frac{3}{4}$	4	3 years
10	$\frac{10}{12} = \frac{5}{6}$	6	5 years
11	$\frac{11}{12}$	12	11 years
12	$\frac{12}{12} = \frac{1}{1}$	1	-
24	$\frac{24}{12} = \frac{2}{1}$	1	-

Periodicity and cycle length of the temporally aggregated process for different values of the aggregation frequency (k). The seasonality in the disaggregate process is set equal to 12.

Table 2: Belgian deficit (real terms). Monthly estimation and quarterly-annual temporal aggregation results

Variables	**Monthly estimated model**	**Quarterly aggregated model**	**Annual aggregated model**
MU	0.7802e-03	0.0023	0.0094
MA1	-0.2159 (lag 1)	-0.0957 (lag 1)	-0.4291 (lag 1)
MA2	-0.4014 (lag 12)	-0.4014 (lag 4)	0.0111 (lag 2)
σ^2_ε or $\sigma^2_{\varepsilon^*}$	4.1931e-05	9.4580e-05	3.2720e-04
Number of observations	252	252	252
Corresponding model	$ARIMA(0,0,1)(0,1,1)_{12}$	$ARIMA(0,0,1)(0,1,1)_4$	$ARIMA(0,1,2)$

Method of monthly estimation: exact maximum likelihood.

Table 3: Belgian deficit (real terms). Direct estimation results

Variables	**Quarterly estimated model**	**Annual estimated model**
MU	0.0024	0.0097
MA1	-0.3885 (lag 1)	-0.7148 (lag 1)
MA2	-0.3494 (lag 4)	-0.1476 (lag 2)
$\sigma^2_{\varepsilon^*}$	9.0723e-05	1.1827e-04
Number of observations	84	21
Corresponding model	$ARIMA(0,0,1)(0,1,1)_4$	$ARIMA(0,1,2)$

Method of estimation: exact maximum likelihood.

Figure 1: Representation of the temporal aggregation mechanism for $k = 12$.

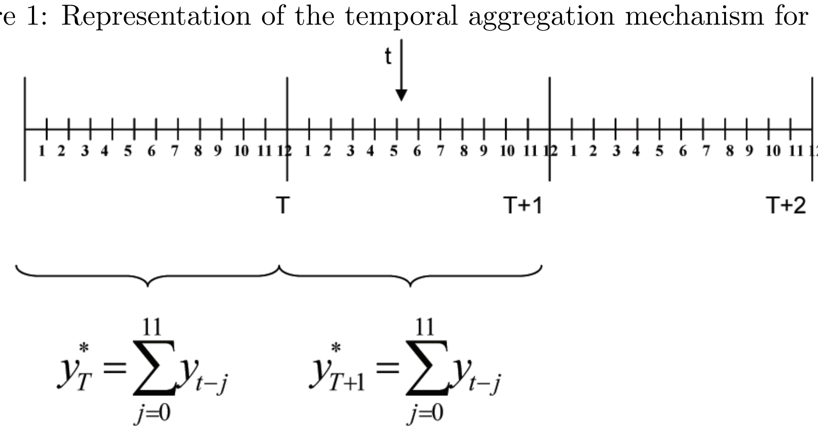


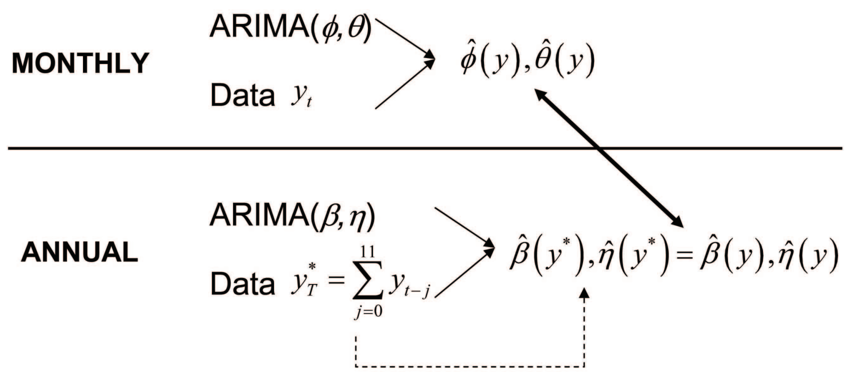
Table 4: Summary of temporal (flow) aggregation - stock aggregation results.

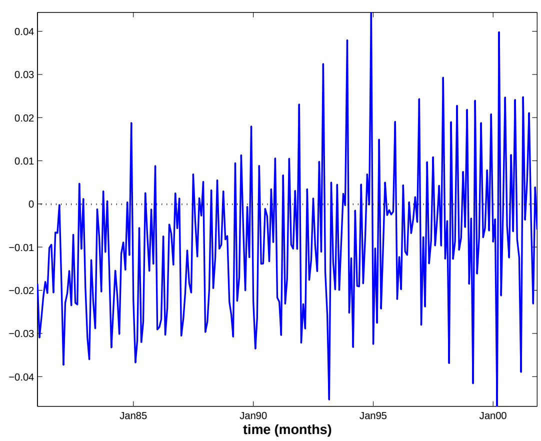
Disaggregate ARIMA model	Aggregate ARIMA model	Aggregate regular AR order	Aggregate seasonal AR order	Aggregate regular MA order	Aggregate seasonal MA order
TEMPORAL (FLOW) AGGREGATION					
$AR(p)$	$ARMA(p,r)$	p	-	$r = \left\lfloor \frac{(p+1)(k-1)}{k} \right\rfloor$	-
$ARMA(p,q)$	$ARMA(p,r)$	p	-	$r = \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor$	-
$ARIMA(p,d,q)$	$ARIMA(p,d,r)$	p	-	$r = \left\lfloor \frac{p(k-1)+(d+1)(k-1)+q}{k} \right\rfloor$	-
$ARIMA(p,d,q)(P,D,Q)_s$	$ARIMA(p,d,r)(P,D,R)_{s^*}$	p	P	$r = \left\lfloor \frac{p(k-1)+(d+1)(k-1)+q}{k} \right\rfloor$	$R = \left\lfloor \frac{(P+D)s^*k+(Q-P-D)s}{k} \right\rfloor$
STOCK AGGREGATION					
$AR(p)$	$ARMA(p,r)$	p	-	$r = \left\lfloor \frac{p(k-1)}{k} \right\rfloor$	-
$ARMA(p,q)$	$ARMA(p,r)$	p	-	$r = \left\lfloor \frac{p(k-1)+q}{k} \right\rfloor$	-
$ARIMA(p,d,q)$	$ARIMA(p,d,r)$	p	-	$r = \left\lfloor \frac{p(k-1)+d(k-1)+q}{k} \right\rfloor$	-
$ARIMA(p,d,q)(P,D,Q)_s$	$ARIMA(p,d,r)(P,D,R)_{s^*}$	p	P	$r = \left\lfloor \frac{p(k-1)+d(k-1)+q}{k} \right\rfloor$	$R = \left\lfloor \frac{(P+D)s^*k+(Q-P-D)s}{k} \right\rfloor$

Disaggregate ARIMA model	Aggregate ARIMA model (flow aggr.)	Aggregate AR order		Aggregate MA order	Lag of aggregate stock exogenous
$ARIMAX(p,d,q)$	$ARIMAX(p,d,r)$	p		r is the greater of $\left\lfloor \frac{(p+d+1)(k-1)+q}{k} \right\rfloor$ and $\left\lfloor \frac{(p+d+1)(k-1)+m-1+\tilde{d}}{k} \right\rfloor + \left\lfloor \frac{(v+\tilde{d})(k-1)+w}{k} \right\rfloor$	$\left\lfloor \frac{(p+d+1)(k-1)+m-1+\tilde{d}}{k} \right\rfloor$

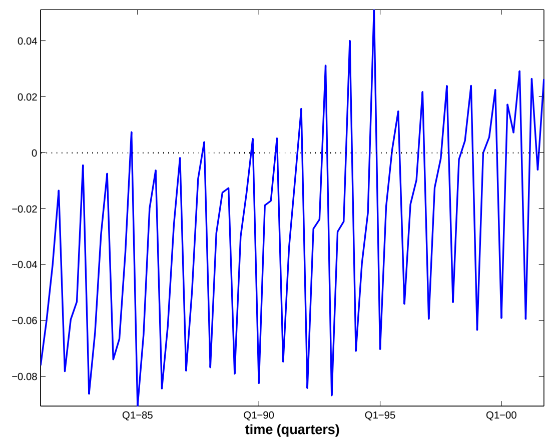
Disaggregate GARCH model	Aggregate GARCH model (flow aggr.)	Aggregate ARMA AR order	Aggregate GARCH AR order	Aggregate ARMA MA order	Aggregate GARCH MA order
$ARMA(p,q)GARCH(P,Q)$	$ARMA(p,r)GARCH(R,R)$	p	$R = \tilde{r} + \frac{1}{2}r(r+1)$	$r = \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor$	$R = \tilde{r} + \frac{1}{2}r(r+1)$

Figure 2: Relation between the parameters of the aggregate (e.g. annual) and disaggregate (e.g. monthly) models.

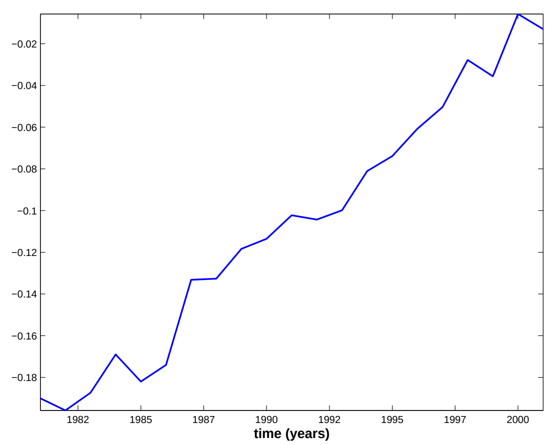




Time series plot of monthly deficit for Belgium (real terms), January 1981-December 2001, y-axis is in billion of Euros.



Time series plot of quarterly deficit for Belgium (real terms), first quarter 1981-last quarter 2001, y-axis is in billion of Euros.



Time series plot of annual deficit for Belgium (real terms), 1981-2001, y-axis is in billion of Euros.

Figure 3: Belgian deficit in real terms at different frequencies.

Département des Sciences Économiques
de l'Université catholique de Louvain
Institut de Recherches Économiques et Sociales

Place Montesquieu, 3
1348 Louvain-la-Neuve, Belgique

ISSN 1379-244X D/2005/3082/044