BANCA D'ITALIA

EUROSISTEMA

Temi di discussione

(Working papers)

Temporal aggregation of univariate and multivariate time series models: A survey

by Andrea Silvestrini and David Veredas

August 2008

Number 685

TEMPORAL AGGREGATION OF UNIVARIATE AND MULTIVARIATE TIME SERIES MODELS: A SURVEY

by Andrea Silvestrini* and David Veredas**

Abstract

We present a unified and up-to-date overview of temporal aggregation techniques for univariate and multivariate time series models explaining in detail how these techniques are employed. Some empirical applications illustrate the main issues.

JEL Classification: C10, C22, C32, C43.
Keywords: temporal aggregation, ARIMA, seasonality, GARCH, vector ARMA, spurious causality, multivariate GARCH.

Contents

---

* Bank of Italy, Economics, Research and International Relations and Université Catholique de Louvain.
** ECARES, Université Libre de Bruxelles and Université Catholique de Louvain, Belgium.

1 Introduction[1]

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

The choice of the frequency clearly 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. For instance, 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 are the annual data a function of the quarterly data, but the annual model is also a function of the quarterly model. Moreover, the quarterly estimated model is richer, information-wise, as the number of observations used for estimation is four times larger than for the annual model.

The way in which these two models interact is the subject of this survey. In a univariate and multivariate times series context, i.e. ARIMA, GARCH, vector ARMA and multivariate 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 studied in econometric literature for the last 35 years, and general conditions have been obtained in terms of order conditions (i.e. polynomial lag length), parameter estimation, asymptotic behaviour, etc. A selected literature consists of Amemiya and Wu (1972), Tiao (1972), Brewer (1973), Wei (1978, 1990), Weiss (1984), Stram and Wei (1986), Lütkepohl (1987), Nijman and Palm (1990), Drost and Nijman (1993), Marcellino (1999), Breitung and Swanson (2002) and Hafner (2004). Yet, to our knowledge, a complete up-to-date survey of the methodology is currently unavailable.

In a nutshell, deriving the low frequency model from the high frequency model involves two stages.

[1]The authors would like to thank Luc Bauwens, Pierluigi Daddi, Christian M. Hafner, Marc Hallin, Theo Nijman, Giacomo Sbrana, Luc Van Meensel, Bas Werker, the ECORE seminar and the Stockholm ENTER 2006 Jamboree participants for helpful comments and suggestions. The views expressed are those of the authors and do not necessarily reflect those of the Bank of Italy. The scientific responsibility is assumed by the authors. Final version forthcoming in: Journal of Economic Surveys (doi: 10.1111/j.1467-6419.2007.00538.x). The definitive version is available at www.blackwell-synergy.com

First, time series models are specified in terms of lag polynomials of a given order. The technique of temporal aggregation allows us 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 questions such as: 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 is it? 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 variables aggregate may take different forms. Two aggregation schemes are often found in economics: *stock* and *flow*. Stock, also called systematic sampling, refers to aggregation in which the aggregated variable is the result of sampling every k periods from the high frequency variable. For instance, annual observations may be obtained by sampling every four periods of quarterly observations. Flow, also called temporal aggregation, refers to aggregation in which the aggregated variable is the sum, every k period, of the high frequency variable. Hence the annual observations are the sum of the quarterly observations every four periods. Rates and indexes, such as interest rate, unemployment rate or CPI, are stock variables while GDP, public deficit or financial returns are examples of flow variables.

The analysis of temporal aggregation starts with the seminal article of Amemiya and Wu (1972). They show that, if the original variable is generated by an AR model of order p, the aggregate variable follows an AR model of order p with MA residuals structure. Tiao (1972) and Amemiya and Wu (1972) study the issue of information loss due to aggregation. They compare - theoretically and with simulations - predictors based on low and high frequency data. They conclude that the optimal predictor built from the high frequency data performs remarkably well with respect to the optimal predictor built from the low frequency sample.[2] Brewer (1973) presents a generalization of the results obtained by Amemiya and Wu for ARMA models with exogenous variables (ARMAX models). Wei (1978) derives the model structure for temporally aggregated data when the high frequency model includes seasonal polynomials. He shows that if the frequency of aggregation is the same as the seasonal frequency (for instance, intra-annual seasonality and annual aggregation), the aggregate model reduces to a model without seasonality. The first author to investigate temporal aggregation for non-stationary models, specifically IMA(d,q) models,

[2]This theoretical result, although correct, is arguable at an empirical level. For instance, Silvestrini *et al.* (2008) show, in an empirical application, that the forecasting capabilities of the temporally aggregated model outperform the ones of the model estimated at the higher frequency. See also Abraham (1982).

is Tiao (1972). Weiss (1984) discusses flow and stock aggregation schemes for ARIMA models. Stram and Wei (1986) focus on the relationship between the autocovariance function of the original disaggregate series and its aggregate counterpart. They also find that, in some special cases, the autoregressive order of ARIMA models can be reduced after temporal aggregation. Drost and Nijman (1993) derive the order conditions for temporally aggregated univariate GARCH models. They show that when the variable is flow, the parameters of the aggregated model depend on the disaggregate 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.

All the contributions so far quoted deal with univariate models. In a general multivariate framework, Lütkepohl (1987) contains a deep analysis of temporal (and contemporaneous) aggregation for VARMA models. It also examines the impact of temporal aggregation on the efficiency of the forecasts. For instance, if the interest is predicting annual inflation, issues such as "Is it better to predict monthly and to aggregate the predictions or rather to predict using annual data?" are addressed. Marcellino (1999) focuses on temporal aggregation of VARMA models and on the effects of temporal aggregation on several time series properties (such as causality, exogeneity, cointegration, unit roots, seasonal unit roots, impulse response functions, trend-cycles decompositions, etc.). Temporal aggregation of multivariate models is complicated by spurious instantaneous causality that may be induced by time aggregation. This phenomenon has also been analysed by Breitung and Swanson (2002) and by Hafner (2004).

Two issues are intrinsically related to temporal aggregation. First, a vast literature analyses the problem 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. We may then be interested in computing the orders and the parameters of the disaggregate (monthly) model starting from the aggregate (annual) one, estimated from data. The whole subject is made more complex by parameter identification issues, see Palm and Nijman (1984). In other words, the same low frequency model may disaggregate to several high frequency models, which are observationally equivalent at the low frequency.

Second, temporal aggregation is not the only kind of aggregation. Many other economic phenomena may be analysed from a cross-section perspective. This scheme of aggregation, through individuals rather than through time, is called *contemporaneous* or *cross-section aggregation*. If these individual series are known to follow a stationary ARMA process, it is possible to investigate whether the aggregate observed series follows an ARMA process as well. We refer the reader to Granger and Morris (1976), Lütkepohl (1984, 1987), Granger (1987, 1990), among others, for a thorough discussion. Contemporaneous aggregation of GARCH models has been analysed by Nijman and Sentana (1996). Meddahi and Renault (2004) study temporal aggregation of square-root stochastic autoregressive volatility models. Zaffaroni (2007) focuses on aggregation of exponential stochastic volatility models and non-linear moving average models.

To conclude, a comment is due on what we do not cover in this survey. Temporal aggregation is a vast field and almost any subject in time series analysis may be investigated within this framework. Non-linearities, long memory, random aggregation, time continuous aggregation, spatial aggregation or factor models are issues that, although fascinating, are not discussed.[3]

The outline of the rest of the paper is as follows: Section 2 introduces the notation, the aggregation schemes and the intuitive foundations of the technique. Section 3 presents temporal aggregation for ARIMA types of models, including an empirical application based on macroeconomic data. Section 4 surveys temporal aggregation of GARCH models, focusing on the GARCH$(1,1)$ and on an empirical application based on financial data. Section 5 deals with temporal aggregation of VARMA models and with spurious causality. Section 6 derives results for temporal aggregation of multivariate GARCH (MGARCH) models and presents an empirical application on financial data. Section 7 summarizes and concludes.

Throughout the survey we favour intuition to technicalities, which can be found in the appropriate references. For instance, the second half of Section 2 is wordy but essential for the understanding of the rest of the survey. Yet, the topic is intrinsically technical - this is a methodological survey - and hence the formulae are unavoidable. The structure of all the sections is very similar. After a brief introduction, we present the main result - such that the reader uniquely interested in applications can skip the rest of the section - that we derive, with more or less detail, afterwards. Last, throughout the survey we focus on the flow aggregation scheme. Results for the stock case can be found in summary Tables 4, 5 and 6 and in the Appendix of Silvestrini and Veredas (2005).

[3]Random aggregation techniques are presented by Jorda and Marcellino (2004). Nonlinearities issues are partly discussed in Granger and Lee (1993) and Proietti (2006). Links with time continuous models may be found in Nelson (1990) and Drost and Werker (1996). An analysis of the consequences of aggregation on long memory processes is in Granger (1980a) and in Tsai and Chan (2005). Spatial aggregation has been studied by Giacomini and Granger (2004) and factor models by Forni *et al.* (2000).

2 Notation and Aggregation Schemes

Let y_t be a random variable observed at *high frequency* t. Sample information for the *low frequency* or *aggregate* random variable is assumed to be available only every kth period $(k, 2k, 3k, \ldots)$, where k, an integer value larger that one, is the aggregation frequency. In general, we define the aggregate 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 of order A in the lag operator L that determines the aggregation scheme. The weights w_j are exactly known. Equation (1) embeds two important aggregation schemes: i) *Flow*: $A = k-1$ and $w_j = 1$, for $j = 0, \ldots, A$. Or $W(L) = 1 + L + \ldots + L^{k-1}$, i.e. aggregation of y_t carried out over k periods. For instance, if y_t is monthly and $k = 3$, we get quarterly sums of the monthly observations. ii) *Stock*: $A = 0$ and $w_0 = 1$. One every k observations is kept, the rest being skipped, i.e. $y_t^* = y_{kt}$. For instance, if the observed time unit t is monthly and $k = 3$, y_t is only observed every third period.

Other important cases are also covered by (1): i) average, $w_j = \frac{1}{k}$, for $j = 0, \ldots, k-1$, and ii) weighted average, $w_j = \frac{\chi_j}{k}$, for $j = 0, \ldots, k-1$, where χ_j are the weights that sum to one. Note that flow, averaging and weighted averaging aggregation schemes are rolling sums. In other words, (1) is computed at every time t, which means a sequence of sums that overlap over $k-1$ periods. However, the aggregate series does not overlap. To indicate the aggregate series we introduce another time scale, T, that runs in kt periods. So that $t = \ldots, 0, 1, 2, \ldots$, while $T = \ldots, 0, k, 2k, \ldots$, as is illustrated in Figure 1 for $k = 12$. Thus, we sub-index the aggregated series by T using the notation $y_T^* = y_{kt}^*$. Flow and stock are the schemes most often found in economics. In the following pages we focus on the flow case, although references to the stock case will be made whenever it enhances the comprehension of the technique. Detailed results for the stock aggregation can be found in the Appendix of Silvestrini and Veredas (2005).

[FIGURE 1 ABOUT HERE]

All the results shown in the next sections rely on the same procedure, which we outline intuitively later in this section. Assume that the disaggregate series, y_t, follows the model

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

where $t = \ldots, 0, 1, 2, \ldots$, $\phi(L)$ and $\theta(L)$ are lag polynomials. Likewise, the temporally aggregated series, y_T^*, follows the model

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

where $T = \ldots, 0, k, 2k, \ldots$, $\beta(B)$ and $\eta(B)$ are aggregate lag polynomials and the operator B is in T time units, running in kt periods. If the aggregate data y_T^* are a function of the disaggregate data y_t - given by (1) - 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). 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 are determined, in turn, by the autocovariance structure of y_t. Once these orders are chosen, we can estimate the parameters as $(\hat{\phi}, \hat{\theta}) = (\hat{\phi}(y), \hat{\theta}(y))$. The expected value of y_T^* is also a function of its own past. However, it differs with respect to y_t as $y_T^* = W(L) y_t$. Therefore the AR and MA aggregate polynomial 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, the estimated parameters, $(\hat{\phi}, \hat{\theta})$, should be such that they incorporate all the maximum information at the minimum cost. 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 can already be extracted. First, we not only aggregate data, we 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: it refers i) to the lag structures of the AR and MA polynomials and ii) to the corresponding parameters. Second, (ϕ, θ) are estimated with all the disaggregate observations. Thus, $\hat{\beta}(\hat{\phi}, \hat{\theta})$ and $\hat{\eta}(\hat{\phi}, \hat{\theta})$ contain all the information of the high frequency sample. This gives a more accurate estimate of the parameters, in terms of consistency and efficiency, than if they were estimated from y_T^*, which has k times fewer observations. 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 situations where decisions are taken, say, annually, but information is available, for instance, monthly. It is not necessary to wait until 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.[4]

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

[4] See Silvestrini *et al.* (2008) for an application to the French public deficit.

the cornerstone of the method. In general, the AR and MA polynomials of the disaggregate model expressed in terms of their roots are multiplied by $T(L)$, i.e. $T(L)\phi(L)y_t = T(L)\theta(L)\varepsilon_t$. The resulting AR polynomial, $T(L)\phi(L)$, has powers of L only divisible by k. We also redefine $L^k = B$. In this way y_t is transformed into y^*_T. Furthermore, the order of the AR polynomial remains the same under temporal aggregation. 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 moving average part of the model is calculated multiplying the disaggregate MA polynomial by the $T(L)$ operator, i.e. $T(L)\theta(L)\varepsilon_t$. The product $T(L)\theta(L)$ includes some AR components, the aggregation scheme (both in $T(L)$) and the MA part. We therefore end up with two MA aggregate polynomials: $T(L)\theta(L)\varepsilon_t$ and $\eta(B)\varepsilon^*_T$. Using deterministic rules, we infer the order of the aggregate MA polynomial, $\eta(B)$, from the order of $T(L)\theta(L)$. Last, we compute the parameters in $\eta(B)$ equating the autocovariance structures of both MA polynomials, $T(L)\phi(L)\varepsilon_t$ and $\eta(B)\varepsilon^*_T$. The result is a non-linear system of equations that can be easily solved.

When y_t is not stationary, there is seasonality or exogenous variables, the conditional variance of the error term has GARCH effects, or for any multivariate extension, 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 is inferred straightforwardly and the MA structure is computed equating the autocovariance structures of the disaggregate and aggregate models.

3 ARIMA Class

In this section we present temporal aggregation for models belonging to the ARIMA class. Section 3.1 deals with pure AR models, 3.2 with ARMA models, 3.3 with ARIMA models, 3.4 with ARIMAX and 3.5 with seasonal ARIMA models. The whole section ends with an empirical application. The main references are Amemiya and Wu (1972), Brewer (1973), Wei (1978), Weiss (1984) and Stram and Wei (1986).

3.1 AR Models

The AR(p) model for y_t is defined as

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

where $\phi(L) = 1 - \phi_1 L - \ldots - \phi_p L^p$ is an autoregressive polynomial of order p and ε_t is a white noise error term with zero mean and constant variance σ_ε^2. Let $\delta_j, j = 1, \ldots, p$, be the distinct inverted roots of $\phi(L)$ polynomial, each assumed to lie inside the unit circle. We express $\phi(L)$ in terms of δ_j as $\phi(L) = \prod_{j=1}^{p}(1 - \delta_j L)$. The following result shows how to derive the appropriate specification for the temporally aggregated variable y_T^*.

Result 1 *The temporal aggregation of y_t as specified in model (4), denoted y_T^*, is represented by an ARMA(p, r) where*

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

and $\lfloor b \rfloor$ indicates the integer part of a real number b.

To derive this result we follow the intuitive guidelines outlined in Section 2. 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 form[5]

$$T(L) = \left[\frac{1 - L^k}{1 - L}\right] \prod_{j=1}^{p} \left[\frac{1 - \delta_j^k L^k}{1 - \delta_j L}\right] \tag{5}$$

which has two parts. The first involves the ratio of two polynomials which equals $\sum_{j=0}^{k-1} L^j$, i.e. the temporal aggregation scheme. The denominator of the product contains the inverted roots of the AR polynomial and its numerator contains the same roots, but powered by the aggregation frequency. Multiplying both sides of (4) by $T(L)$ we get

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

where y_t^* is the temporally aggregated variable with temporal index t operating on the disaggregate time unit T. The powers of the product $T(L)\phi(L)$ 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

[5]The general form of this polynomial will be given when we introduce seasonality. To begin, we start with the easiest form, adding terms as we augment the model.

order is unchanged by temporal aggregation:[6]

$$\beta(B)y_T^* = \prod_{j=1}^{p} \left[1 - \delta_j^k L^k\right] y_T^* \tag{7}$$

We now focus on the right-hand side (RHS) of (6) which represents a moving average structure

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

with $\delta_0 = 0$. This is a linear combination of aggregated error terms. More precisely, $T(L)\varepsilon_t$ is a moving average of order $(p+1)(k-1)$. It is expressed in terms of t rather than T, however. To switch the time frequency and to get the appropriate order of $\eta(B)$, we divide the order of $T(L)\theta(L)$ by k. That is, the order of $\eta(B)$ corresponds to $\lfloor k^{-1}(p+1)(k-1)\rfloor$ in aggregate time units T. Amemiya and Wu (1972) prove that the MA polynomial for the aggregate series y_T^* is invertible, i.e. the MA roots lie outside the unit circle.

Finally, we infer the parameters of the temporally aggregated model. For the AR part it is trivial, as shown 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 they are MA polynomials, only the variance and the first r autocovariances at time T are different from zero. As the number of unknowns is also $r+1$, the problem consists in solving a non-linear system of equations with as many unknowns as equations.

3.2 ARMA Models

Next, we augment (4) assuming that y_t follows an autoregressive moving average (ARMA) model

$$\phi(L)y_t = \theta(L)\varepsilon_t \tag{8}$$

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. These polynomials are assumed to have their roots outside the unit circle and to have no common roots.

Until now we have introduced the polynomial $T(L)$ as the cornerstone of the method. However, we have not justified it. Why does it take the form it does and not another one? Why do the aggregate AR

[6]We impose that temporally aggregated ARIMA type models display no hidden periodicity of order k. This means that in (6) the AR and MA polynomials share no roots such that no root cancellation occurs and no AR order reduction is observed. We refer to Stram and Wei (1986) for further discussion.

parameters appear so easily while the MA are more time consuming to calculate? Brewer (1973) provides a nice interpretation of $T(L)$ that we summarize later in the section.

To obtain the temporally aggregated model each side of (8) has to be multiplied by $T(L)$. 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 (1973) in matrix form as

$$\begin{bmatrix} 1 & 0 & \dots & 0 \\ t_1 & 1 & \dots & 0 \\ t_2 & t_1 & 1 & 0 \\ \vdots & \vdots & \vdots & \vdots \\ t_p & t_{p-1} & \dots & 1 \\ t_h & t_{h-1} & \dots & t_{h-p} \\ 0 & t_h & \dots & \dots \\ 0 & 0 & t_h & \dots \\ 0 & 0 & 0 & t_h \end{bmatrix} \begin{bmatrix} 1 \\ -\phi_1 \\ \vdots \\ -\phi_p \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \beta_1 \\ \beta_1 \\ \vdots \\ \beta_2 \\ \beta_2 \\ \vdots \\ \beta_c \end{bmatrix} \tag{9}$$

where the matrices are of order $(p+h+1) \times (p+1)$, $(p+1) \times 1$ and $(c+1)k \times 1$, respectively, and they stand for $T(L)$, $\phi(L)$ and $\beta(L)$. The only dimension that we know is that of $\phi(L)$, $(p+1) \times 1$. The dimensions of $T(L)$ and $\beta(L)$ depend on h and c respectively, which are unknown. However, we may recover them through several order restrictions that are implicit in (9). The product $T(L)\phi(L)$ has the dimension $(p+h+1) \times 1$. A first restriction that helps us to recover h and c is that the number of rows on both sides of (9) are identical:

$$p+h+1 = (c+1)k \Rightarrow p+h+1 = ck+k \tag{10}$$

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

$$h = (c+1)(k-1) \Rightarrow h = ck+k-c-1 \tag{11}$$

Substituting (10) in (11) we get $h = p+h+1-c-1 \Rightarrow c = p$. That is, the order of the AR polynomial is unchanged under temporal aggregation ($p = c$). 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 of order kr is

$$E[T(L)\theta(L)\varepsilon_t T(L)\theta(L)\varepsilon_{t-kr}]$$

which is different from zero for $r \leq \lfloor k^{-1}\left((p+1)(k-1)+q\right)\rfloor$. Thus, the maximum aggregate MA order is $r=\lfloor k^{-1}\left((p+1)(k-1)+q\right)\rfloor$. This derivation brings the following result.

Result 2 *The temporal aggregation of y_t as specified in model (8), denoted y_T^*, is represented by an $ARMA(p,r)$ where*

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

Indeed, multiplying both sides of (8) by $T(L)$,

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

with $\theta_0=1$. The AR part of the model is treated exactly in the same way as in the pure autoregressive case. And the MA part is of order $(p+1)(k-1)+q$ in time t or of order $\lfloor k^{-1}\left((p+1)(k-1)+q\right)\rfloor$ in time T.

3.3 ARIMA Models

If we make the further assumption that y_t follows an integrated ARMA (ARIMA), equation (8) becomes

$$\phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t \tag{12}$$

where d is a real integer denoting the order of integration. 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:

$$\bar{T}(L)=\left[\frac{1-L^k}{1-L}\right]^{d+1}\prod_{j=1}^{p}\left[\frac{1-\delta_j^k L^k}{1-\delta_j L}\right]. \tag{13}$$

Note the difference with respect to (5). The first component is now powered by $d+1$ instead of 1. Multiplying the disaggregate model by $\bar{T}(L)$, the high frequency unit roots vanish and only $(1-L^k)^d$ is left, i.e. a polynomial operating on T with d unit roots. The following result stems from Weiss (1984).

Result 3 *The temporal aggregation of y_t as specified in model (12), denoted y_T^*, is represented by an ARIMA(p,d,r) where*

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

Multiplying each side of (12) by the augmented $\bar{T}(L)$ operator (13) yields

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

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

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

where ξ_l is a function of δ_j $(j=1,\ldots,p)$ and θ_j $(j=1,\ldots,q)$. The order of the RHS polynomial is $r = k^{-1}\lfloor p(k-1)+(d+1)(k-1)+q \rfloor$ in time T.

Last, note that the roots of the AR polynomial in the temporally aggregated model are the kth powers of the AR roots in the disaggregate model. As explained by Tiao (1972), for large values of the sampling interval k the AR aggregate coefficients decrease in size, that is, the AR terms disappear and the MA part starts to be dominated by the unit roots, leading to an IMA(d,d) model. See Rossana and Seater (1995) for an empirical investigation on the effects of an increase in the aggregation frequency.

3.4 ARIMAX Models

In econometric models it is very often the case that an economic variable is not only explained by itself but also by explanatory variables, assumed to be exogenous. In this section we study temporal aggregation when these explanatory variables are present. The main reference is Brewer (1973). For the sake of simplicity, we work with only one non-seasonal exogenous variable. Extensions to more than one variable are straightforward.

Two important clarifications are in order. First, earlier we made the distinction between different types of aggregation schemes, stock and flow being 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. Four possibilities are at hand, but in order to introduce some variety we assume that

y is flow, as usual, while x is stock.[7] Second, the new twist in complexity comes from 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 aggregated model are a function of the orders of the model for x. This also implies that we get not only one equation that defines the order of the aggregate MA polynomial (as we have done until now), but also one for the aggregate component of x.

The ARIMAX$(p,d,q)(m)$ model for y_t is defined as

$$\phi(L)(1-L)^d y_t = \theta(L)\varepsilon_t + C(L)(1-L)^{\tilde{d}} x_t \tag{14}$$

where x_t is a stock exogenous variable and $C(L)$ its associated polynomial of length m. This exogenous variable follows an ARIMA$(v,\tilde{d},w)$,

$$D(L)(1-L)^{\tilde{d}} x_t = F(L)u_t \tag{15}$$

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. All the polynomials are assumed to have their roots outside the unit circle and to have no common roots.

The problem consists in performing temporal aggregation of the ARIMAX model (14) with a stock exogenous variable obeying (15). The following result gives the order conditions and the number of past lags of x in the temporally aggregated model. The derivation is slightly different from the previous ones because of the presence of x. Indeed, due to the time aggregation mechanism some of the x in the aggregated specification are observations at lags not divisible by the aggregation frequency. In other words, the aggregated model for y_T^* includes lags of x between T and $T-k$. This problem is circumvented expressing x terms as a function of their own aggregated past and their error terms using their own model in (15). Furthermore, this implies that in the temporally aggregated model at time t we have two error terms, ε_t and u_t. The MA order of the aggregated model is the greater of the two.

Result 4 *Given the ARIMAX$(p,d,q)(m)$ model in (14) and the stock exogenous variable following (15), the temporally aggregated series of y_t, denoted y_T^*, is represented by an ARIMAX(p,d,r)(a) where r is the greater of*

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

[7] This is the case that we study in detail, but Table 5 shows the order conditions for all possible combinations.

with

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

3.5 Seasonal ARIMA Models

When aggregating seasonal cycles we may distinguish three cases: when the aggregation frequency is smaller than the cycle (for example, we aggregate from monthly to quarterly and seasonality is annual), when it is the same (for example, we aggregate from monthly to annual) and when it is larger (for example, we aggregate from monthly to biannual). In the first case, the aggregated model still has some seasonal components. In the last two cases seasonality vanishes. In the third case, in particular, all the seasonal components become a regular AR(1) and a seasonal unit root becomes a regular unit root.

The ARIMA$(p,d,q) \times (P,D,Q)_s$ for y_t is defined as

$$\phi(L)\Phi(L^s)(1-L)^d(1-L^s)^D y_t = \theta(L)\Theta(L^s)\varepsilon_t \tag{16}$$

where $\Phi(L^s) = 1 - \ldots - \Phi_P L^{Ps}$ and $\Theta(L^s) = 1 - \ldots - \Theta_Q L^{Qs}$ are polynomials in the seasonal lag operator of length Ps and Qs, s is the seasonal frequency and D the number of unit roots in seasonality. In addition, let $\tau_1^s, \ldots, \tau_P^s$ be the inverted roots of $\Phi(L^s)$ such that $\Phi(L^s) = \prod_{i=1}^{P} [1-(\tau_i L)^s]$. The polynomials are assumed to have their roots outside the unit circle and to have no common roots.

To account for seasonality, we introduce a new operator

$$A(L) = \left[\frac{1-L^{ks^*}}{1-L^s}\right]^D \prod_{i=1}^{P} \left[\frac{1-(\tau_i L)^{ks^*}}{1-(\tau_i L)^s}\right] \tag{17}$$

Like $\bar{T}(L)$ it consists of two parts. The first accounts for the seasonal unit roots, while the second includes the inverted roots of the seasonal autoregressive polynomial. However, $A(L)$ does not include the aggregation scheme, while $\bar{T}(L)$ does. Another difference with $\bar{T}(L)$ is given by s^*. It is the seasonal frequency of the temporally aggregated process, which may take anomalous values without 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 every 3 periods (i.e. quarters), $s^* = 4$, which has a very clear economic meaning. However, if we aggregate every five periods, s^* equals 60 or 5 years, that is, the temporally aggregated process takes five years to end a cycle in the same month as the original process. The reason for 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 years to have a 5-month period that ends in a month which is a multiple of 12. Two interesting cases are when k equals 12 and 24. For $k = 12$, the aggregation frequency equals the high frequency seasonal frequency, $k = s$. In this case seasonality vanishes. For $k = 24$, seasonality also vanishes. This drives us to some important conclusions. First, if $k < s$, there is still some seasonality in the temporally aggregated process. Second, if k is a multiple of s, the seasonal cycle remains constant. Last, if k is equal or larger than s, seasonality vanishes.

Result 5 *The temporal aggregation of y_t as specified in model (16), denoted y_T^*, is represented by an $ARIMA(p,d,r) \times (P,D,R)_{s^*}$ where*

$$r = \left\lfloor \frac{(p+1)(k-1) + d(k-1) + q}{k} \right\rfloor \quad \textit{and} \quad R = \left\lfloor \frac{(P+D)s^*k + (Q-P-D)s}{k} \right\rfloor$$

Multiplying both sides of (16) by $\bar{T}(L)A(L)$,

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

After some algebra the LHS becomes

$$\prod_{j=1}^{p}[1-\delta_j^k L^k]\prod_{i=1}^{P}[1-(\tau_i L)^{ks^*}](1-L^k)^d(1-L^{ks^*})^D y_t^*$$

i.e. an ARI$(p,d)\times(P,D)_{s^*}$. The RHS equals $\bar{S}(L)A(L)\theta(L)\Theta(L^s)\varepsilon_t^*$, where

$$\bar{S}(L) = \prod_{j=1}^{p}\left[\frac{1-\delta_j^k L^k}{1-\delta_j L}\right]\left[\frac{1-L^k}{1-L}\right]^d \tag{18}$$

and the order of the polynomial in ε_t satisfies the equalities

$$rk = d(k-1) + p(k-1) + (k-1) + q \quad \text{and} \quad Rk = (P+D)s^*k + (Q-P-D)s.$$

Thus y_T^* follows an ARIMA$(p,d,r) \times (P,D,R)_{s^*}$. Note that whenever $P = D = Q = 0$ (no seasonality) R equals zero. These derivations are slightly different from those that may be found in Weiss (1984), since he does not compute the maximum order R.

3.6 Empirical Application to Belgian Federal Deficit

In this section we present an empirical application of the techniques surveyed so far using Belgian public deficit data. We use the *net balance to be financed* - federal deficit in short. This is the definition given by the Belgian Federal Public Service Finance. The federal deficit series is used by the federal Treasury in order to monitor debt management. It is limited to (federal government) Treasury operations. As a consequence, it does not take into account the 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.[8] The data set used contains 252 monthly observations. Our source is the National Bank of Belgium.

Figure 2 shows the Belgian cash deficit series, in real terms, at different frequencies. The top left panel displays the series at monthly frequency, the top right-hand panel at quarterly frequency, while the bottom panel presents the annual series. The monthly and quarterly series have a clear seasonal pattern due to the intra-annual instalments corresponding to tax collection and payments. For instance, advance payments by companies are received in April, July, October and December. As a result, during these months the federal deficit is generally positive. As these instalments are intra-annual, the annual series does not reflect any seasonal pattern. On the basis of the explanations given in Section 3.5, the annual seasonality vanishes because $s = k = 12$. In this regard, note 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 evident that the annual series has a unit root. As we shall see, the monthly series has a seasonal unit root. This nicely dovetails 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 2 ABOUT HERE]

We now perform several estimation exercises. We start estimating a monthly model,[9] then we aggregate it quarterly and annually.[10] To make a comparison we directly estimate the same quarterly and annual models from the quarterly and annual data sets. The monthly series can be represented by an

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

[9] 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 Criterion.

[10] 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/.

ARIMA$(0,0,1)\times(0,1,1)_{12}$ model with intercept

$$(1-L^{12})y_t=\hat{\mu}+(1+\hat{\theta}L)(1+\hat{\Theta}L^{12})\varepsilon_t \tag{19}$$

First, we aggregate the monthly model to quarterly frequency using $T(L)=\sum_{i=0}^{2}L^j$,

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

Letting $B=L^3$ operate in time T, this quarterly aggregated model is an ARIMA$(0,0,1)\times(0,1,1)_4$,

$$\left(1-B^4\right)y_T^*=M+\left(1+\eta_1 B\right)\left(1+E_1B^4\right)\varepsilon_T^*$$

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

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

Note that everything on the RHS here is known, as it has already been estimated. The system has four equations with three unknowns and can be easily solved.[11]

[TABLE 1 ABOUT HERE]

Second, we perform temporal aggregation into annual frequency. To do this, we multiply both sides of (19) by $T(L)=\sum_{i=0}^{11}L^j$.

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

If we let $B=L^{12}$ operate in time $T=1,2,\ldots$, the aggregate model is an ARIMA$(0,1,2)$,

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

[11]Note that the third equation is nested in the fourth, and is therefore redundant. It yields a system of three equations with three unknowns that are exactly identified.

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

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

We also estimate the quarterly and annual models directly from the aggregate observations (quarterly and annual data sets). All the results are shown in Table 1. Some conclusions may be drawn. We divide them into *within*, meaning the analysis of the monthly and temporally aggregated models, and *between*, meaning the differences between the models in the upper and bottom panels of Table 1.

Within conclusions: first, the intercepts 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 possess seasonal MA components and it becomes an MA(2). Moreover, even if we aggregate 12 periods, the temporally aggregated 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 is reflected in Figure 2, 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 intercepts are very similar, meaning that there is a coherence and the aggregated models behave properly. Second, by contrast, the MA parameters differ, especially for the annual frequency. A possible cause is that the number of annual observations is 21, compared with the 252 of the aggregated model, suggesting merely that asymptotic properties apply. 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, direct estimation and temporal aggregation produce comparable results.

4 ARMA-GARCH Models

In this section we review the main results on temporal aggregation of univariate ARMA-GARCH models, which 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 possible. The main reference is Drost and Nijman (1993).

The ARMA(p, q)-GARCH(P, Q) model for y_t is defined as

$$\begin{aligned} \phi(L) y_t &= \theta(L)\varepsilon_t, \quad \varepsilon_t = \xi_t \sqrt{h_t}, \quad \xi_t \sim D(0,1) \\ b(L) h_t &= \psi + a(L)\varepsilon_t^2 \end{aligned} \tag{20}$$

where ε_t is a sequence of stationary errors with zero mean, variance h_t and finite fourth moments.[12] Furthermore, $\phi(L) = 1 - \sum_{i=1}^{p} \phi_i L^i$, $\theta(L) = \sum_{i=0}^{q} \theta_i L^i$, $a(L) = \sum_{i=1}^{Q} a_i L^i$ and $b(L) = 1 - \sum_{i=1}^{P} b_i L^i$. The polynomials $b(L)$ and $b(L) - a(L)$ are assumed to have roots outside the unit circle and to be invertible.[13] Drost and Nijman (1993) introduce three definitions of GARCH models (strong, semi-strong and weak) and show that only the weak one is *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 requires that errors are built as a martingale difference sequence, i.e. with conditional mean equal to zero and variance equal to the GARCH model. Weak GARCH is characterized by defining h_t as the best linear predictor of ε_t^2 in terms of a constant, lagged values of ε_t^2 and lagged values of ε_t. Consequently, differences between expected and realized first and second moments of the error term are uncorrelated.[14]

Drost and Nijman (1993) show that only *weak* GARCH models are *closed* under temporal aggregation. If, for instance, daily returns follow a weak GARCH process, then weekly and monthly returns will

[12] $D(0, 1)$ represents a generic distribution with zero mean and unit variance.

[13] A necessary and sufficient condition for the second-order stationarity of the GARCH(P, Q) model is $\Sigma_{i=1}^{P} b_i + \Sigma_{i=1}^{Q} a_i < 1$. This condition, due to Bollerslev (1986), is binding for all the GARCH models presented in the paper. Moreover, as reported by Li *et al.* (2002), $\psi > 0$, $b_i \geq 0$, $a_i \geq 0$ $(\forall i)$ are sufficient conditions to guarantee that $h_t > 0$. Less restrictive conditions are given by Drost and Nijman (1993, p. 911). Concerning the ARMA part, we assume that the roots of $\phi(L)$ and $\theta(L)$ polynomials lie outside the unit circle and that no common roots are present.

[14] This third definition will become clearer shortly.

follow a weak GARCH as well, with corresponding parameter adjustments. Expressed differently, when aggregating from y_t to y_T^*, the distribution function may change or ε_T^* may not be a martingale difference sequence anymore (semi-strong GARCH definition). But differences between realized and expected first two moments remain uncorrelated (weak GARCH definition). This idea is precisely defined as follows.

Definition 1 *The variable ε_t follows a symmetric weak GARCH model if ε_t is uncorrelated and if ψ, $a(L)$ and $b(L)$ can be chosen such that*

$$P[\varepsilon_t|\varepsilon_{t-1},\varepsilon_{t-2},\ldots]=0 \quad \textit{and} \quad P[\varepsilon_t^2|1,\varepsilon_{t-1},\varepsilon_{t-2},\ldots,\varepsilon_{t-1}^2,\varepsilon_{t-2}^2,\ldots]=h_t,$$

where $P[x_t|1,\varepsilon_{t-1},\varepsilon_{t-2},\ldots,\varepsilon_{t-1}^2,\varepsilon_{t-2}^2,\ldots]$ 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,

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

Therefore, a weak GARCH model is defined as h_t being the projection on a constant term, on lagged ε_t^2 and lagged ε_t variables.[15] Moreover, the expected difference between the realized and the projected value is required to be orthogonal, up to the second moment, to the realized value. The following result shows the orders of the aggregated model.

Result 6 *The temporal aggregation of y_t as specified in (20), denoted y_T^*, is represented by an ARMA(p,r) with weak GARCH(R,R) errors where*

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

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

The derivation of the general case, not shown here, is structured in two steps. The first step consists in inferring the aggregate conditional mean (i.e. the ARMA part of the model) from the high frequency one. The second step deals with the conditional variance. The GARCH model is defined on the variance of the residuals and hence the treatment is slightly different from the conditional mean. The starting point is rewriting the GARCH model as an ARMA model. Then, following the logic of the previous sections,

[15] Hafner (2004) shows that defining h_t as the projection on a constant, lagged terms of ε_t^2 *and* lagged terms of ε_t is not a necessary condition to ensure closeness of the model under temporal aggregation.

an operator is applied to the transformed model. However, this operator differs from all the others as the GARCH model is not on y_T^* but on ε_T^{*2}. Moreover, the square of ε_t is to be aggregated, instead of ε_t itself. In the flow case, this creates cross-products and the fourth moment structure turns out to be relevant to determining the properties of the temporally aggregated model.

As an example, consider a GARCH$(1,1)$ model, the most used volatility model. For the sake of exposition we focus on an aggregation frequency $k = 2$. We can write the temporally aggregated GARCH model in ARMA form at time t as

$$(\varepsilon_t + \varepsilon_{t-1})^2 = 2(1 + a + b)\psi + (a + b)^2(\varepsilon_{t-2} + \varepsilon_{t-3})^2 + w_t$$

where $w_t = \iota_t + (1+a)\iota_{t-1} + (a - b(a+b))\iota_{t-2} - b(a+b)\iota_{t-3} + 2\varepsilon_t\varepsilon_{t-1} - 2(a+b)^2\varepsilon_{t-2}\varepsilon_{t-3}$ and $\iota_t = \varepsilon_t^2 - h_t$. Since $E(w_t w_{t-2j}) = 0$, $\forall j > 1$, the aggregated model is a GARCH$(1,1)$

$$h_T^* = \psi^* + b^* h_{T-1}^* + a^*(\varepsilon_{T-1}^*)^2$$

where $\psi^* = (1 + a + b)\psi$, $a^* = (a + b)^2 - b^*$ and b^* is the solution of

$$\frac{b^*}{1 + (b^*)^2} = \frac{E(w_t w_{t-2})}{E((w_t)^2)} = \frac{g(a, b, \kappa_y)(a + b)^2 - l(a, b)}{g(a, b, \kappa_y)(1 + (a + b)^4) - 2l(a, b)}$$

where

$$\begin{aligned}
g(a, b, \kappa_y) &= 2(1 - b)^2 + 4\frac{(1 - a - b)^2(1 - b^2 - 2ab)}{(\kappa_y - 1)(1 - (a + b)^2)} + 4\frac{(1 - 2(a + b) + (a + b)^2)(a - ab(a + b))}{1 - (a + b)^2} \\
l(a, b) &= (a - ab(a + b))\frac{1 - (a + b)^4}{1 - (a + b)^2} \\
\kappa_y &= \kappa_\xi \frac{1 - (a + b)^2}{1 - (a + b)^2 - (\kappa_\xi - 1)a^2}
\end{aligned}$$

and κ_ξ is the unconditional kurtosis of the rescaled innovation, i.e. $\xi_t = \varepsilon_t/\sqrt{h_t}$. Last, the aggregated unconditional kurtosis is

$$\kappa_{y^*} = 3 + \frac{\kappa_y - 3}{2} + 6(\kappa_y - 1)\frac{1 - 2(a + b) + (a + b)^2(a - ab(a + b))}{4(1 - a - b)^2(1 - b^2 - 2ab)}$$

Similar expressions for any k may be found in Drost and Nijman (1993, p. 916).

4.1 Empirical Application to MSCI Index

We apply temporal aggregation results for a GARCH$(1, 1)$ model to the daily MSCI - Morgan Stanley Corporate Index - of Indonesia from April 1, 2001 to April 30, 2006. The data set consists of 1303 observations. The top panel of Figure 3 plots the series of returns at different aggregation frequencies: daily, weekly, biweekly, monthly. At the daily frequency, at least two outliers are clearly visible and a remarkable volatility clustering is smoothed out as the aggregation frequency increases.

[FIGURE 3 ABOUT HERE]

We estimate at daily frequency a GARCH$(1, 1)$ model with Student-t errors with ν degrees of freedom. The corresponding unconditional kurtosis of the rescaled innovations is $\frac{3\nu-6}{\nu-4}$ for $\nu > 4$.

[TABLE 2 ABOUT HERE]

Estimation results are provided in Table 2. The stationarity condition $(\hat{a} + \hat{b} < 1)$, essential for aggregation, is satisfied. With these estimated parameters we derive the GARCH$(1, 1)$ parameters in the case of flow aggregation. The aggregation frequency ranges from 2 to 50. The bottom panels of Figure 3 display the aggregated parameters, ψ^*, a^*, b^* and κ_y^*, at different aggregation frequencies. The intercept, ψ^*, increases monotonically with k. The shock parameter, a^*, decreases very quickly for small aggregation frequencies and then decays at a slower rate, approaching zero when $k = 50$. This is expected, since, as derived by Drost and Nijman (1993, p. 916), $a^* = (a + b)^k - b^*$. The persistence parameter, b^*, decreases sharply to zero. This explains why the volatility clustering tends to disappear as the aggregation frequency increases. However, the light increase for small values of k is noticeable. Finally, κ_{y^*} starts out at around 9 and slowly decays towards 6. Note that with an aggregation frequency of $k = 50$, κ_{y^*} is clearly above 6.4. Therefore, the rate of convergence towards Gaussianity is rather slow.

5 Causality and VARMA Models

Multivariate models commonly feature properties under temporal aggregation not observed when dealing with univariate models. Marcellino (1999), for instance, analyses the effects of temporal aggregation on exogeneity, causality, cointegration, unit roots, seasonal unit roots, impulse response functions and trend-cycles decompositions. He finds that cointegration and unit roots are invariant to temporal aggregation, whereas exogeneity, causality, seasonal unit roots, impulse response functions and trend-cycles decompositions are not. In fact, a rich literature focuses on the effects of temporal aggregation on causality. Once observations are temporally aggregated, the observed causal structure may be different from the original one. Quoting Weiss (1984, p. 280), "Some care needs to be taken in causality testing, as causality is defined for the true processes and not for the equation on the (temporally) aggregated or sampled data". Earlier studies on this topic are Quenouille (1957), Sims (1971), Tiao and Wei (1976) and Geweke (1978). Wei (1982) shows that temporal aggregation may convert a one-way causality into bidirectional causality. Christiano and Eichenbaum (1987) consider the temporal aggregation bias leading to spurious Granger causality. This issue is empirically investigated in the context of causal relationships between growth rate of money and aggregate output.

Granger (1969) formally introduces the concept of *spurious instantaneous causality*, meaning instantaneous causality between variables observed at the low frequency without any causality at the high frequency. See also Pierce and Haugh (1977) and Granger (1980b). Granger (1988) studies in detail some possible explanations for this phenomenon, one being temporal aggregation. Indeed, spurious instantaneous causality may be found whenever the interval at which data are collected is lower than the frequency at which data are supposed to be generated, the so-called *natural frequency*. Therefore, spurious instantaneous causality and temporal aggregation are tightly linked. Renault *et al.* (1998) discuss spurious instantaneous causality distinguishing between the true component of the observed causality in a continuous time model and the spurious component, produced by model discretization. More recently, Breitung and Swanson (2002) examine the impact of temporal aggregation on instantaneous causality within the context of VAR models.

To explain how instantaneous causality between temporally aggregated time series may arise even if there is no causality between disaggregated time series we need two related concepts: Granger causality and instantaneous causality. Both notions are based on optimal linear forecasts, although several alternative definitions have been proposed. The seminal paper by Granger (1969) addresses the question of whether an economic variable can help forecast another economic variable, giving rise to the general concept of *Granger causality*. This article has stimulated an independent body of applied literature through the years. One of the earliest examples is Sims (1972), which provides some operational novelties in the analysis of the causal ordering for money and income. Technical definitions of Granger causality and instantaneous causality may be found in the already cited Granger (1969), Pierce and Haugh (1977) and Granger (1980b), among many others. More formal work on causality for linear processes is conducted by Florens and Mouchart (1985). Lütkepohl (2005) derives useful characterizations of Granger causality and instantaneous causality in the context of VAR models. We refer to all these contributions and the references therein for deeper insights and formal expressions. Hereafter we provide some intuition, avoiding technical details.

When contemporaneous values of y_t are useful for forecasting contemporaneous values of x_t, we say that y_t *instantaneously causes* x_t. Pierce and Haugh (1977) suggest to analyse the cross-correlations between the innovations of the series to detect instantaneous causality. Instantaneous cross-correlation is indeed crucial for instantaneous causality. In particular, it may be proved that a necessary condition for linear processes is that their correlation is not null, i.e. $\rho(y_t, x_t) \neq 0$. Spurious instantaneous causality between x_t and y_t occurs when the variables do not possess any Granger causal relationship at the natural frequency, but for the aggregated variables we find that $\rho(x_t, y_t) \neq 0$. Consider, for instance, the following example - taken from Bramati (2005):[16]

$$
\begin{aligned}
x_t &= a_1 x_{t-1} + a_2 y_{t-1} + \varepsilon_t^x \\
y_t &= a_4 y_{t-1} + \varepsilon_t^y
\end{aligned}
$$

It is assumed that y_t Granger causes x_t with one lag period (indeed $a_2 \neq 0$). Furthermore, by hypothesis there is no instantaneous causality between the variables at the specified frequency, i.e. $\rho(y_t, x_t) = 0$. We sample the two variables every two periods ($k = 2$), obtaining

$$
\begin{aligned}
x_{2t} &= a_1^2 x_{t-2} + (a_1 + a_4) a_2 y_{t-2} + a_1 \varepsilon_{t-1}^x + a_2 \varepsilon_{t-1}^y + \varepsilon_t^x \\
y_{2t} &= a_4^2 y_{t-2} + a_4 \varepsilon_{t-1}^y + \varepsilon_t^y
\end{aligned}
$$

Here we observe that y_{2t} Granger causes x_{2t} with one aggregate lag period (indeed $(a_1 + a_4)a_2 \neq 0$). Moreover, note that ε_{t-1}^y term is present in both equations. Therefore, it turns out that $\rho(y_{2t}, x_{2t}) \neq 0$. Hence, a misleading instantaneous causality has been induced by the aggregation procedure. This example confirms that, as pointed out by Marcellino (1999), causality is *not* a time series property *invariant* to temporal aggregation. However, there are conditions that can be imposed to rule out spurious instantaneous causality. Breitung and Swanson (2002) give sufficient conditions to ensure that,

[16] A similar example may be found in Gulasekaran and Abeysinghe (2002).

as the temporal aggregation frequency goes to infinity, there is no instantaneous causality among the aggregated variables (proposition 3, p. 655).

We now turn to the effects of temporal aggregation on model structure. We assume that $\mathbf{y}_t$ is an n-variate time series. Each $\mathbf{y}_t$ is a column vector whose components are denoted $y_{t,1},\ldots,y_{t,n}$. The n-variate VARMA(p,q) model for $\mathbf{y}_t$ may be expressed in compact form as

$$\mathbf{G}(L)\mathbf{y}_t = \mathbf{M}(L)\boldsymbol{\varepsilon}_t, \tag{21}$$

where $\mathbf{G}(L) = \mathbf{I}_n - \mathbf{G}_1 L - \ldots - \mathbf{G}_p L^p$ and $\mathbf{M}(L) = \mathbf{I}_n + \mathbf{M}_1 L + \ldots + \mathbf{M}_q L^q$ are $n \times n$ matrices of lag polynomials, $(\mathbf{G}_i)_{i=1}^p$ and $(\mathbf{M}_i)_{i=1}^q$ are $n \times n$ parameter matrices, $\boldsymbol{\varepsilon}_t$ is a white noise error vector with zero mean and non-singular positive-definite variance matrix $\boldsymbol{\Sigma}_\varepsilon$, i.e. $\boldsymbol{\varepsilon}_t \sim WN(\mathbf{0}, \boldsymbol{\Sigma}_\varepsilon)$.

A comprehensive analysis of temporal (and contemporaneous) aggregation in a general multivariate framework is in Lütkepohl (1987), which brings into focus the impact of temporal and contemporaneous aggregations on the efficiency of the forecasts. Temporally aggregation of VARMA models is also discussed, based on the commonly named macro processes. Marcellino (1999) derives the generation mechanism of a temporally aggregated process, assuming that the disaggregated one is an integrated VARMA (VARIMA) model. Most concepts and findings reviewed in this section are based upon this last contribution.

The starting point is model (21) for the disaggregate variable $\mathbf{y}_t$. The aim is to provide a complete characterization (in terms of orders and parameters) of the corresponding temporally aggregated model for $\mathbf{y}_T^*$. The link between the models for $\mathbf{y}_t^*$ and $\mathbf{y}_T^*$ is, once again, given by the polynomial matrix $\tilde{\mathbf{T}}(L)$. The mechanism is similar to that already outlined in the univariate case. We pre-multiply (21) by $\tilde{\mathbf{T}}(L)$

$$\tilde{\mathbf{T}}(L)\mathbf{G}(L)\mathbf{y}_t^* = \tilde{\mathbf{T}}(L)\mathbf{M}(L)\boldsymbol{\varepsilon}_t^* \tag{22}$$

We let $\bar{\mathbf{G}}(L) = \tilde{\mathbf{T}}(L)\mathbf{G}(L)$ and $\mathbf{N}(L) = \tilde{\mathbf{T}}(L)\mathbf{M}(L)$. Hence $\bar{\mathbf{G}}(L)\mathbf{y}_t^* = \mathbf{N}(L)\boldsymbol{\varepsilon}_t^*$. The degree of $\tilde{\mathbf{T}}(L)$ is $p(k-1)$ and therefore $\tilde{\mathbf{T}}(L)\mathbf{G}(L)$ has degree $pk - p + p = pk$, i.e. k times the original autoregressive lag length. In addition, the coefficients in $\tilde{\mathbf{T}}(L)\mathbf{G}(L)$ that are not a multiple of L^k must be null as a consequence of the temporal aggregation scheme. We define the vectors of matrices $\mathbf{G}^v$ and $\tilde{\mathbf{T}}^v$ as

$$\mathbf{G}^v = \left[\underset{n\times n}{\mathbf{G}_1}, \quad \mathbf{G}_2, \quad \ldots, \quad \mathbf{G}_p \;, \quad \mathbf{0}, \quad \ldots \quad \mathbf{0}\; \right], \tilde{\mathbf{T}}^v = \left[\underset{n\times n}{\tilde{\mathbf{T}}_1}, \tilde{\mathbf{T}}_2, \quad \ldots, \quad \tilde{\mathbf{T}}_{pk-p} \;\right]$$

of size $n \times npk$ and $n \times np(k-1)$, respectively. Furthermore, let the matrix $\mathbf{G}^m$ be

$$\mathbf{G}^m = \begin{bmatrix} -\mathbf{I}_n & \mathbf{G}_1 & \dots & \mathbf{G}_p & \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & -\mathbf{I}_n & \dots & \mathbf{G}_{p-1} & \mathbf{G}_p & \mathbf{0} & \dots & \mathbf{0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{0} & \mathbf{0} & \mathbf{0} & \dots & \mathbf{G}_p \end{bmatrix}$$

of size $np(k-1) \times npk$. In a similar way, $\mathbf{G}^v_{-k}$ and $\mathbf{G}^m_{-k}$ are a $n \times np(k-1)$ and $np(k-1) \times np(k-1)$ matrices obtained by deleting the k^{th} columns of $\mathbf{G}^v$ and $\mathbf{G}^m$. The following result provides the orders of the VARMA(p, q) model for $\mathbf{y}_t$.

Result 7 *If $\left|\mathbf{G}^m_{-k}\right| \neq 0$, the temporal aggregation of $\mathbf{y}_t$ as specified in (21), denoted $\mathbf{y}^*_T$, is represented by a VARMA(p, r) model*

$$\bar{\mathbf{G}}(B)\mathbf{y}^*_T = \bar{\mathbf{M}}(B)\boldsymbol{\varepsilon}^*_T \tag{23}$$

where $\bar{\mathbf{G}}(B)$ and $\bar{\mathbf{M}}(B)$ are the AR and MA polynomial matrices of order p and r, with

$$\begin{aligned} r = p - s \quad & \textit{for} \quad sk < p - q + 1 \leq (s+1)k, \quad s = 0, 1, \dots, p \\ r = p \quad & \textit{for} \quad \quad p = q \\ r = p + 1 + s \quad & \textit{for} \quad sk < q - 1 - p < (s+1)k, \quad s = 0, 1, \dots \end{aligned}$$

The aggregate white noise error vector is $\boldsymbol{\varepsilon}^*_T \sim WN\left(\mathbf{0}, \boldsymbol{\Sigma}_{\varepsilon^*}\right)$.[17] The coefficients of the aggregate AR polynomial matrix $\bar{\mathbf{G}}(B) = \mathbf{I}_n - \bar{\mathbf{G}}_1 B - \dots - \bar{\mathbf{G}}_p B^p$ are the k^{th} columns of the matrix product $\mathbf{G}^v_{-k}(\mathbf{G}^m_{-k})^{-1}\mathbf{G}^m - \mathbf{G}^v$.

The coefficients of the aggregate MA polynomial matrix $\bar{\mathbf{M}}(B) = \mathbf{I}_n + \bar{\mathbf{M}}_1 B + \dots + \bar{\mathbf{M}}_r B^r$ and the aggregate variance matrix $\boldsymbol{\Sigma}_{\varepsilon^*}$ can be recovered as a solution of the non-linear system

$$\begin{aligned} \sum_{i=0}^{r} \bar{\mathbf{M}}_i \boldsymbol{\Sigma}^*_{\varepsilon} \bar{\mathbf{M}}'_i &= \sum_{i=0}^{pk-p+q} \mathbf{N}_i \boldsymbol{\Sigma}_{\varepsilon} \mathbf{N}'_i \\ \bar{\mathbf{M}}_j \boldsymbol{\Sigma}_{\varepsilon^*} + \sum_{i=1}^{r-j} \bar{\mathbf{M}}_{i+j} \boldsymbol{\Sigma}_{\varepsilon^*} \bar{\mathbf{M}}'_i &= \mathbf{N}_{jk} \boldsymbol{\Sigma}_{\varepsilon} + \sum_{i=1}^{pk-p+q-jk} \mathbf{N}_{i+jk} \boldsymbol{\Sigma}_{\varepsilon} \mathbf{N}'_i \qquad (j = 1, \dots, r) \end{aligned} \tag{24}$$

[17] Upper bounds for the orders of the autoregressive and moving average polynomial matrices may be also found in the monograph by Lütkepohl (1987), but they refer to a VARMA model (for the temporally aggregated variable) in final equations form, i.e. a VARMA in which the AR polynomial matrix is expressed as $\bar{\mathbf{G}}(B) = \bar{g}(B)\mathbf{I}_n$, where $\bar{g}(B) = 1 - \bar{g}_1 B - \dots - \bar{g}_p B^p$ is a scalar operator and $\bar{g}_p \neq 0$. Additional explanations are given in Sections 6.1 and 6.5 of the same book.

whenever the invertibility condition $\left|\bar{\mathbf{M}}(z)\right| \neq 0$, $z \leq 1$, is satisfied and $\bar{\mathbf{M}}(0) = \mathbf{I}_n$. The analogy with the univariate case is clear. Indeed, this is an extension of a simpler method proposed by Marcellino (1996) to determine the coefficients of temporally aggregated ARIMA models. In the univariate case results coincide with those reported by Brewer (1973) and Weiss (1984).

The matrix coefficient of L^i in $\tilde{\mathbf{T}}(L)\mathbf{G}(L)$ on the LHS of (22) corresponds to the i^{th} matrix column of $\tilde{\mathbf{T}}^v\mathbf{G}^m - \mathbf{G}^v$. Since the matrix coefficients of $\bar{\mathbf{G}}(B)$ in (23) are those multiple of L^k in $\tilde{\mathbf{T}}^v\mathbf{G}^m - \mathbf{G}^v$, those not associated with a multiple of L^k are restrained to be zero (as already explained, this is a direct consequence of temporal aggregation). For that reason, the matrix coefficients in $\tilde{\mathbf{T}}(L)$ must be chosen in such a way as to satisfy the system of equations

$$\tilde{\mathbf{T}}^v\mathbf{G}^m_{-k} - \mathbf{G}^v_{-k} = \mathbf{0}$$

If $\left|\mathbf{G}^m_{-k}\right| \neq 0$, this system has a unique solution: $\tilde{\mathbf{T}}^v = \mathbf{G}^v_{-k}(\mathbf{G}^m_{-k})^{-1}$. Therefore, the coefficients in the AR polynomial matrix $\bar{\mathbf{G}}(B)$ are the k^{th} columns of the matrix product $\mathbf{G}^v_{-k}(\mathbf{G}^m_{-k})^{-1}\mathbf{G}^m - \mathbf{G}^v$, as stated in the result. This is an equivalent result to the univariate case.

Concerning the MA polynomial matrix $\bar{\mathbf{M}}(B)$ in (23), the non-linear system in (24) is built equating the autocovariance structures of $\bar{\mathbf{M}}(B)\boldsymbol{\varepsilon}^*_T$ and of $\tilde{\mathbf{T}}(L)\mathbf{M}(L)\boldsymbol{\varepsilon}^*_t$. As is well known, two models with the same autocovariance function are the same model. Indeed, these two differently parameterized models are the result of the same temporal aggregation scheme applied to (21). Imposing the equivalence of the autocovariance functions, we are able to recover the unknown coefficients of the MA polynomial matrix and the unknown variance matrix $\mathbf{\Sigma}_{\varepsilon^*}$.

6 Multivariate GARCH Models

The multivariate GARCH literature is extensive, including, among others, Bollerslev *et al.* (1988), Engle *et al.* (1990), Engle and Kroner (1995), Kroner and Ng (1998), Hafner and Herwartz (1998), van der Weide (2002), Tse and Tsui (2002), Engle (2002) and Kawakatsu (2006). Estimation issues are discussed by Gouriéroux (1997), Jeantheau (1998), van der Weide (2002) and Comte and Lieberman (2003). Causality in volatility is studied by Comte and Lieberman (2000). Fourth moment properties are investigated by Hafner (2003). Bauwens *et al.* (2006) provide a review of the whole MGARCH literature, including model specifications and inference methods. Gouriéroux (1997), Campbell *et al.* (1997), Franses and van Dijk

(2000) and Lütkepohl (2005) are books that include multivariate GARCH models. Contributions dealing with temporal aggregation of MGARCH models are by far less numerous. The performance of quasi maximum likelihood and non-linear least squares estimation methods applied to temporally aggregated GARCH models is tested by Hafner and Rombouts (2003).

Most of the theoretical results are, however, derived by Hafner (2004). This paper, which constitutes our main reference, addresses several interesting questions: closeness of weak multivariate GARCH models, multivariate volatility forecasting, multivariate realized volatility, estimation issues and spurious instantaneous causality in temporally aggregated MGARCH models. Necessary conditions for spurious instantaneous causality (in volatility) are given, one being zero conditional covariance between two series. Financial time series, however, tend to be correlated at high frequencies. For that reason, instantaneous causality (in volatility) is often observed also at the high frequency, not only at the low frequency. And it is not spuriously created by temporal aggregation. Empirically, therefore, spurious instantaneous causality for MGARCH models is a much less relevant feature compared with spurious instantaneous causality for VARMA.

In what follows we focus on the *weak* version of MGARCH models. As shown by Hafner (2004), the class of weak MGARCH models is the only one to be closed under temporal aggregation. Strong or semi-strong MGARCH models are not closed. This result constitutes an analogue to the univariate case. Moreover, we refer to MGARCH models in vector specification form (VEC), following the original definition given by Bollerslev *et al.* (1988). This is the most general *linear* specification. It nests the diagonal VEC (DVEC) model of Bollerslev *et al.* (1988), the BEKK model of Engle and Kroner (1995) and the factor GARCH (F-GARCH) model of Engle *et al.* (1990). However, it does not nest the constant conditional correlation (CCC) model of Bollerslev (1990) and the dynamic conditional correlation (DCC) model of Engle (2002) and Tse and Tsui (2002).[18] Without loss of generality, we consider temporal aggregation of a weak MGARCH$(1,1)$. As pointed out by Bauwens *et al.* (2006), orders higher than $(1,1)$ are rarely encountered in empirical applications.

Let $\mathbf{H_t}$ be a positive-definite symmetric matrix of dimension $n \times n$. The N-variate multivariate GARCH$(1,1)$ model for ε_t may be specified in VEC form as

$$\underset{N\times 1}{\mathbf{h}_t} = \underset{N\times 1}{\boldsymbol{\psi}} + \mathbf{A} \underset{N\times 1}{\boldsymbol{\eta}_{t-1}} + \mathbf{B} \underset{N\times 1}{\mathbf{h}_{t-1}} \tag{25}$$

[18] These are both non-linear specifications of MGARCH models.

where $\mathbf{h}_t = \text{vech}(\mathbf{H}_t)$, $\boldsymbol{\eta}_t = \text{vech}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t')$, $\boldsymbol{\psi} = \text{vech}(\boldsymbol{\Omega})$ and $\mathbf{A}$, $\mathbf{B}$, $\boldsymbol{\Omega}$ are $N \times N$ parameter matrices, with $N = n(n+1)/2$. If $\mathbf{h}_t$ is assumed to represent the best linear predictor of $\boldsymbol{\eta}_t$ in terms of a constant and lagged values of $\boldsymbol{\eta}_t$, then $\boldsymbol{\varepsilon}_t$ follows a *weak* MGARCH$(1,1)$. The MGARCH$(1,1)$ may be expressed as a VARMA$(1,1)$ for $\boldsymbol{\eta}_t$

$$\boldsymbol{\eta}_t = \boldsymbol{\psi} + \mathbf{Q}\,\boldsymbol{\eta}_{t-1} - \mathbf{B}\,\boldsymbol{\iota}_{t-1} + \boldsymbol{\iota}_t \tag{26}$$

where $\boldsymbol{\iota}_t = \boldsymbol{\eta}_t - \mathbf{h}_t$, $\mathbf{Q} = \mathbf{A} + \mathbf{B}$. All the eigenvalues of the matrix $\mathbf{Q}$ have modulus smaller than one. If $\boldsymbol{\varepsilon}_t$ follows a weak MGARCH$(1,1)$ and possesses the fourth moments finite, then $\boldsymbol{\iota}_t$ is a weak vector white noise with positive-definite variance matrix $\boldsymbol{\Sigma}_\iota = E[\boldsymbol{\iota}_t\boldsymbol{\iota}_t'] < \infty$, $E[\boldsymbol{\iota}_t\boldsymbol{\iota}_s'] = \mathbf{0}$ $(\forall t \neq s)$. Consequently, (26) is a *weak* VARMA$(1,1)$ representation for $\boldsymbol{\eta}_t$. From (26) we obtain a VMA(∞) representation for $\boldsymbol{\eta}_t$

$$\boldsymbol{\eta}_t = \boldsymbol{\sigma} + \sum_{l=0}^{\infty} \tilde{\mathbf{B}}_l \boldsymbol{\iota}_{t-l}$$

where $\sum_{l=0}^{\infty} \tilde{\mathbf{B}}_l \boldsymbol{\iota}_{t-l} = (\mathbf{I}_N - \mathbf{Q})^{-1}(\mathbf{I}_N - \mathbf{B})\,\boldsymbol{\iota}_t$ is an infinite MA polynomial matrix (of order N) and $\boldsymbol{\sigma} = (\mathbf{I}_N - \mathbf{Q})^{-1}\boldsymbol{\psi}$ is a constant term. Note that $\boldsymbol{\sigma}$ coincides with the vectorized unconditional covariance matrix of $\boldsymbol{\varepsilon}_t$, i.e. $\text{vech}(\boldsymbol{\Sigma}_\varepsilon)$, which exists and is finite if and only if all the eigenvalues of the matrix $\mathbf{Q}$ have modulus smaller than one. The VMA(∞) representation is useful since it allows us to recover, in a straightforward way, the autocovariance structure of $\boldsymbol{\eta}_t$. The first two unconditional moments are: $E[\boldsymbol{\eta}_t] = \boldsymbol{\sigma}$ and $\boldsymbol{\Sigma}_\eta = E[\boldsymbol{\eta}_t\boldsymbol{\eta}_t'] = \sum_{l=0}^{\infty} \tilde{\mathbf{B}}_l \boldsymbol{\Sigma}_\iota \tilde{\mathbf{B}}_l'$. Note that $\boldsymbol{\Sigma}_\eta$ is a matrix of fourth moments of $\boldsymbol{\varepsilon}_t$. The autocovariance matrix is

$$\boldsymbol{\Gamma}_\eta(\tau) = E((\boldsymbol{\eta}_t - \boldsymbol{\sigma})(\boldsymbol{\eta}_{t-\tau} - \boldsymbol{\sigma})') = \sum_{l=0}^{\infty} \tilde{\mathbf{B}}_{\tau+l} \boldsymbol{\Sigma}_\iota \tilde{\mathbf{B}}_l'$$

Let us now focus on the temporal aggregation mechanism. We consider, without loss of generality, $k = 2$. The usual notation indicates the partial sums $\boldsymbol{\varepsilon}_t^* = \boldsymbol{\varepsilon}_t + \boldsymbol{\varepsilon}_{t-1}$, bearing in mind that each element of the sum is a vector of dimension $n \times 1$. As in the univariate case, temporal aggregation of MGARCH models is influenced by the fourth moment structure. The aggregated model, indeed, is not expressed in terms of $\boldsymbol{\varepsilon}_t$, but in terms of $\boldsymbol{\eta}_t = \text{vech}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t')$.[19] Consider the vector $\boldsymbol{\eta}_t^* = \text{vech}(\boldsymbol{\varepsilon}_t^*\boldsymbol{\varepsilon}_t^{*'})$ of dimension $N(N+1)/2 \times 1$. It is compounded by the sum of the squares and cross-products of the aggregated

[19]This is similar to the univariate case, where ε_t and not ε_t^2 is aggregated.

process $\boldsymbol{\varepsilon}_t^*$

$$\boldsymbol{\eta}_t^* = \text{vech}\begin{bmatrix} (\varepsilon_{t,1}^*)^2 & \varepsilon_{t,1}^*\varepsilon_{t,2}^* & \dots & \varepsilon_{t,1}^*\varepsilon_{t,N}^* \\ \vdots & (\varepsilon_{t,2}^*)^2 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{t,N}^*\varepsilon_{t,1}^* & \dots & \dots & (\varepsilon_{t,N}^*)^2 \end{bmatrix} \tag{27}$$

Each of the starred vectors corresponds to a partial sum of vectors as a result of temporal aggregation. We can further develop (27)

$$\boldsymbol{\eta}_t^* = \begin{bmatrix} \varepsilon_{t,1}^2 + \varepsilon_{t-1,1}^2 + (2\varepsilon_{t,1}\varepsilon_{t-1,1}) \\ \vdots \\ \varepsilon_{t,N}\varepsilon_{t,1} + \varepsilon_{t-1,N}\varepsilon_{t-1,1} + (\varepsilon_{t,N}\varepsilon_{t-1,1} + \varepsilon_{t-1,N}\varepsilon_{t,1}) \\ \varepsilon_{t,2}^2 + \varepsilon_{t-1,2}^2 + (2\varepsilon_{t,2}\varepsilon_{t-1,2}) \\ \vdots \\ \varepsilon_{t,N}\varepsilon_{t,2} + \varepsilon_{t-1,N}\varepsilon_{t-1,2} + (\varepsilon_{t,N}\varepsilon_{t-1,2} + \varepsilon_{t-1,N}\varepsilon_{t,2}) \\ \varepsilon_{t,3}^2 + \varepsilon_{t-1,3}^2 + (2\varepsilon_{t,3}\varepsilon_{t-1,3}) \\ \vdots \\ \varepsilon_{t,N}^2 + \varepsilon_{t-1,N}^2 + (2\varepsilon_{t,N}\varepsilon_{t-1,N}) \end{bmatrix}$$

We observe that $\boldsymbol{\eta}_t^*$ is the sum of contemporaneous squared terms and non-contemporaneous cross-products (the ones in parentheses). Consequently

$$\boldsymbol{\eta}_t^* = \text{vech}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_t') + \text{vech}(\boldsymbol{\varepsilon}_{t-1}\boldsymbol{\varepsilon}_{t-1}') + \mathbf{v}_t^* = \boldsymbol{\eta}_t + \boldsymbol{\eta}_{t-1} + \mathbf{v}_t^* \tag{28}$$

where $\mathbf{v}_t^*$ represents the non-contemporaneous cross-products, while $\boldsymbol{\eta}_t$ and $\boldsymbol{\eta}_{t-1}$ are contemporaneous squared terms: $\mathbf{v}_t^* = 2\mathbf{D}_n^+(\text{vech}(\boldsymbol{\varepsilon}_t\boldsymbol{\varepsilon}_{t-1}'))$, where $\mathbf{D}_n^+ = (\mathbf{D}_n'\mathbf{D}_n)^{-1}\mathbf{D}_n'$ and $\mathbf{D}_n$ is the duplication matrix.

We define the variance matrix of $\mathbf{v}_t^*$ as $\boldsymbol{\Sigma}_{v^*} = E[\mathbf{v}_t^*\mathbf{v}_t^{*'}]$. Each term in $\mathbf{v}_t^*$ has zero mean and is uncorrelated with any other term of $\mathbf{v}_t^*$ and with $\boldsymbol{\eta}_{t-j}$, $j = 0, 1, \dots, k-1$. The term $\mathbf{v}_t^*$ acts as a noise term that is added to the sum of the high frequency second order process $\boldsymbol{\eta}_t$. Uncorrelation of $\mathbf{v}_t^*$ is a crucial property to obtain closeness of the temporally aggregated MGARCH. The MGARCH model expressed at the low frequency, indeed, needs to possess a VARMA representation in which the (low frequency) error term is a weak white noise. Since $\mathbf{v}_t^*$ is incorporated in the (low frequency) VARMA error term, we need $\mathbf{v}_t^*$ to be uncorrelated. We hereafter state the main result of this section, which is valid for any aggregation frequency k.

Result 8 *The temporal aggregation of* $\boldsymbol{\eta}_t$ *in (26), denoted* $\boldsymbol{\eta}_T^*$*, follows a VARMA*$(1,1)$ *model*

$$\boldsymbol{\eta}_T^* \quad = \quad \boldsymbol{\psi}^* + (\mathbf{A}^* + \mathbf{B}^*)\boldsymbol{\eta}_{T-1}^* + \mathbf{w}_T^* \tag{29}$$

where $\mathbf{w}_T^* = \boldsymbol{\iota}_T^* - \mathbf{B}^*\boldsymbol{\iota}_{T-1}^*$*. Moreover,* $\boldsymbol{\iota}_T^*$ *is a weak white noise with variance matrix* $\boldsymbol{\Sigma}_{\boldsymbol{\iota}^*} = E[\boldsymbol{\iota}_T^*\boldsymbol{\iota}_T^{*'}]$*.*

Furthermore, the temporal aggregation of $\mathbf{h}_t$ *in (25), denoted* $\mathbf{h}_T^*$*, follows a weak MGARCH*$(1,1)$ *model*

$$\mathbf{h}_T^* = \boldsymbol{\psi}^* + \mathbf{A}^*\boldsymbol{\eta}_{T-1} + \mathbf{B}^*\mathbf{h}_{T-1}^*$$

where $\boldsymbol{\psi}^* = k(\boldsymbol{I}_N + (\mathbf{A}+\mathbf{B}) + \ldots + (\mathbf{A}+\mathbf{B})^{k-1})\boldsymbol{\psi}$ *and the* $\mathbf{B}^*$ *matrix can be recovered as a solution of the non-linear system*

$$\mathbf{B}^*\boldsymbol{\Gamma}_{w^*}(1)\mathbf{B}^{*'} + \mathbf{B}^*\boldsymbol{\Sigma}_{w^*} + \boldsymbol{\Gamma}_{w^*}(1) = \mathbf{0} \tag{30}$$

with all the eigenvalues of $\mathbf{B}^*$ *smaller than one in modulus. Once* $\mathbf{B}^*$ *is determined, it is used to calculate* $\mathbf{A}^* = (\mathbf{A}+\mathbf{B})^k - \mathbf{B}^*$*.*

The matrices $\boldsymbol{\Sigma}_{w^*} = E[\mathbf{w}_T^*\mathbf{w}_T^{*'}]$ and $\boldsymbol{\Gamma}_{w^*}(1) = E[\mathbf{w}_T^*\mathbf{w}_{T-1}^{*'}]$ represent the variance and the first-order autocovariance of the aggregate MA(1) vector process in (29). Higher orders autocovariances are null. Expressions for $\mathbf{w}_T^*$, $\boldsymbol{\Sigma}_{w^*}$ and $\boldsymbol{\Gamma}_{w^*}(1)$ and the proof of the result are derived by Hafner (2004). In general, $\boldsymbol{\Sigma}_{w^*}$ and $\boldsymbol{\Gamma}_{w^*}(1)$ matrices are functions of the variance matrix of $\boldsymbol{\iota}_t$, $\boldsymbol{\Sigma}_\iota$, and of the variance matrix of $\mathbf{v}_t^*$, $\boldsymbol{\Sigma}_{v^*}$, i.e. $\boldsymbol{\Sigma}_{w^*} = f(\boldsymbol{\Sigma}_\iota, \mathbf{A}, \mathbf{B})$ and $\boldsymbol{\Gamma}_{w^*}(1) = g(\boldsymbol{\Sigma}_\iota, \mathbf{A}, \mathbf{B})$. The matrices $\boldsymbol{\Sigma}_\iota$, $\mathbf{A}$ and $\mathbf{B}$ may be used to determine $\boldsymbol{\Sigma}_{w^*}$ and $\boldsymbol{\Gamma}_{w^*}(1)$. These are plugged inside (30), which has to be solved numerically. In this way, starting from the disaggregate parameters, i.e. $\boldsymbol{\Sigma}_\iota$, $\mathbf{A}$, $\mathbf{B}$, $\boldsymbol{\psi}$, it is possible to recover the corresponding aggregate parameters, i.e. $\boldsymbol{\Sigma}_{\iota^*}$, $\mathbf{A}^*$, $\mathbf{B}^*$, $\boldsymbol{\psi}^*$.

6.1 Empirical Application to DAX and CAC 40 Indexes

As an illustration, we estimate a bivariate GARCH$(1,1)$ model on two daily log returns series for DAX and CAC 40 indexes. The DAX is the most commonly cited index for measuring returns on the Frankfurt Stock Exchange. It is comprised of the 30 largest stocks traded on the exchange. The CAC 40 consists of the 40 stocks that are most representative of the various economic sectors quoted on the Eurolist market operated by Euronext Paris. Data range from March 2, 1994 to August 25, 2000. The data set consists of 1,693 observations. Top panels of Figure 4 display the returns. Kurtosis statistics (5.54 and 4.72 for DAX and CAC 40, respectively) confirm, as expected, the tail thickness often found in financial time series, as well as the volatility clustering that can be observed in the panels.

[FIGURE 4 ABOUT HERE]

The estimated model is a bivariate diagonal BEKK$(1,1,1)$,[20]

$$
\begin{aligned}
\underset{(2\times 1)}{\mathbf{y}_t} &= \boldsymbol{\mu}_t + \boldsymbol{\varepsilon}_t, \\
\underset{(2\times 1)}{\boldsymbol{\varepsilon}_t} &= \mathbf{H}_t^{1/2}\boldsymbol{\xi}_t, \qquad \boldsymbol{\xi}_t \sim i.i.d.\ \ N(\mathbf{0}, \mathbf{I}_2) \\
\underset{(2\times 2)}{\mathbf{H}_t} &= \tilde{\mathbf{C}}\tilde{\mathbf{C}}' + \tilde{\mathbf{A}}'\boldsymbol{\epsilon}_{t-1}\boldsymbol{\epsilon}_{t-1}'\tilde{\mathbf{A}} + \tilde{\mathbf{B}}'\mathbf{H}_{t-1}\tilde{\mathbf{B}}
\end{aligned} \tag{31}
$$

where $\tilde{\mathbf{C}}$ is a lower triangular parameter matrix and $\mathbf{H}_t^{1/2}$ is a positive-definite parameter matrix such that $\mathbf{H}_t$ is the conditional covariance matrix of $\mathbf{y}_t$. We assume that the disaggregate process $\boldsymbol{\varepsilon}_t$ is strong multivariate GARCH with Gaussian innovations (i.e. whose distribution belongs to the class of spherical distributions with finite fourth moments). Note that a diagonal BEKK$(1,1,1)$ is the easiest specification in terms of number of parameters to estimate, since the parameter matrices $\tilde{\mathbf{A}}$ and $\tilde{\mathbf{B}}$ are required to be diagonal. Estimation results for the bivariate BEKK$(1,1,1)$ model in (31) are given in Table 3. Estimates are in line with results found in the literature. The intercepts are small, though significantly different from zero, and the volatility shows persistence.

[TABLE 3 ABOUT HERE]

As explained by Engle and Kroner (1995), the BEKK$(1,1,1)$ in (31) is equivalent to a VEC$(1,1)$ model in which the corresponding parameter matrices are determined according to

$$
\begin{aligned}
\boldsymbol{\psi} &= \mathrm{vech}(\tilde{\mathbf{C}}\tilde{\mathbf{C}}') \\
\mathbf{A} &= \mathbf{D}_2^{+}(\tilde{\mathbf{A}} \otimes \tilde{\mathbf{A}})'\mathbf{D}_2 \\
\mathbf{B} &= \mathbf{D}_2^{+}(\tilde{\mathbf{B}} \otimes \tilde{\mathbf{B}})'\mathbf{D}_2
\end{aligned} \tag{32}
$$

where $\mathbf{D}_2$ is the duplication matrix of order two and $\mathbf{D}_2^{+}$ is its generalized inverse, i.e. $\mathbf{D}_2^{+} = (\mathbf{D}_2'\mathbf{D}_2)^{-1}\mathbf{D}_2'$. Based on (32), therefore, the estimated BEKK$(1,1,1)$ is equivalent to the following estimated VEC$(1,1)$

[20] This is a different formulation than the VEC one. In the latter, conditions may be imposed on the parameters to ensure that the conditional covariance matrices are positive-definite. See Bauwens *et al.* (2006) for further details. To guarantee the positivity of $\mathbf{H}_t$ without imposing restrictions on the parameters, Engle and Kroner (1995) introduce an alternative parameterization for $\mathbf{H}_t$, known as BEKK. We refer to Engle and Kroner (1995) for additional details and explanations.

model:

$$\text{vech}(\mathbf{H}_\text{t}) = \begin{bmatrix} h_{t,11} \\ h_{t,21} \\ h_{t,22} \end{bmatrix} = \boldsymbol{\psi} + \mathbf{A} \begin{bmatrix} \eta_{t-1,1} \\ \eta_{t-1,2} \\ \eta_{t-1,3} \end{bmatrix} + \mathbf{B} \begin{bmatrix} h_{t-1,11} \\ h_{t-1,21} \\ h_{t-1,22} \end{bmatrix} \tag{33}$$

where

$$\boldsymbol{\psi} = 10^{-5} \begin{bmatrix} 0.1880 \\ 0.1124 \\ 0.2283 \end{bmatrix}, \; \mathbf{A} = \begin{bmatrix} 0.0555 & 0 & 0 \\ 0 & 0.0612 & 0 \\ 0 & 0 & 0.0673 \end{bmatrix}, \; \mathbf{B} = \begin{bmatrix} 0.9316 & 0 & 0 \\ 0 & 0.9244 & 0 \\ 0 & 0 & 0.9172 \end{bmatrix}$$

Bauwens *et al.* (2006) discuss stationarity conditions for multivariate GARCH. For the VEC$(1,1)$ in (33), the maximum eigenvalue of the matrix $(\mathbf{A} + \mathbf{B})$ is smaller than one, hence the overall estimated model is covariance-stationary. Furthermore, "A correct understanding of the fourth moment structure turns out to be essential for the study of temporal aggregation", as explained by Hafner (2004, p. 9). Specifically, it is necessary to check that fourth moments of $\boldsymbol{\varepsilon}_t$ are finite.[21] We rely on Theorem 2 of Hafner (2003, p. 31), where the following expression for the fourth moments of $\boldsymbol{\epsilon}_t$ is provided

$$\text{vech}(\boldsymbol{\Sigma}_\eta) = \mathbf{G}_2 \left(\mathbf{I}_9 - \mathbf{Z}\right)^{-1} \text{vec}(\boldsymbol{\sigma}\boldsymbol{\sigma}')$$

with $\mathbf{G}_2 = c(2(\mathbf{L}^+ \otimes \mathbf{D}_2^+)(\mathbf{I}_2 \otimes \mathbf{C}_{22} \otimes \mathbf{I}_2)(\mathbf{D}_2 \otimes \mathbf{D}_2) + \mathbf{I}_9)$ and $c = 1$, since we are dealing with multivariate Gaussian innovations. Here $\mathbf{L}^+$ and $\mathbf{C}_{22}$ are the elimination and commutation matrices of order 2, respectively. For VEC $(1,1)$ models with spherical innovations, according to Theorem 3 of Hafner (2003, p. 31), we need all the eigenvalues of the matrix

$$\mathbf{Z} = ((\mathbf{A} \otimes \mathbf{A})\mathbf{G}_2 + \mathbf{A} \otimes \mathbf{B} + \mathbf{B} \otimes \mathbf{A} + \mathbf{B} \otimes \mathbf{B}) \tag{34}$$

to have modulus smaller than one for the fourth moments to be finite. The maximum eigenvalue of the $\mathbf{Z}$ matrix in (34) is indeed smaller than one, hence this condition is satisfied.[22] Since the estimated model is covariance-stationary and possesses finite fourth moments, we can proceed to apply results on temporal aggregation of MGARCH models.

The main issue to be addressed is how the model changes when the DAX and CAC 40 log returns series are temporally aggregated, i.e. the behaviour of the elements of the $\mathbf{A}$ and $\mathbf{B}$ parameter matrices

[21] To this end, we remind that $\boldsymbol{\Sigma}_\eta$ is a matrix of such fourth moments.

[22] All the results and the MATLAB codes are available from the authors upon request.

as a function of the aggregation frequency k.[23]

The bottom panels of Figure 4 display the behaviour of the elements of the $\mathbf{B}^*$ and $\mathbf{A}^*$ parameter matrices for k=2, 3, ..., 100. For each of the 99 values of k, the non-linear system in (30) is solved. For any k, once $\mathbf{B}^*$ is determined it is used to calculate $\mathbf{A}^* = (\mathbf{A} + \mathbf{B})^k - \mathbf{B}^*$. For large k, the aggregated parameters converge towards zero as expected. In general, they all tend to decrease as k increases, except for some strange erratic behaviour for aggregation frequencies between 12 and 30. In fact, theorem 2 in Hafner (2004) states that conditional heteroskedasticity of the temporally aggregated process $\boldsymbol{\varepsilon}_t^* = \boldsymbol{\varepsilon}_t + \boldsymbol{\varepsilon}_{t-1} + \ldots + \boldsymbol{\varepsilon}_{t-k+1}$ disappears asymptotically (for $k \rightarrow \infty$). In summary, the main message conveyed is that convergence to zero of the $\mathbf{B}^*$ and $\mathbf{A}^*$ matrices is very slow. This slow decay may be due to the fact the disaggregate process in (33) is close to the stationarity boundary. It could be interesting to assess, empirically, how the convergence changes with different model specifications at the disaggregate frequency. This is out of the scope of the survey and is left for future research.

7 Conclusion

We provide a comprehensive and up-to-date survey of temporal aggregation for univariate and multivariate mean and variance time series models, which has so far been lacking in the literature. We review results for temporal aggregation of AR, ARIMA, ARIMA with seasonality, ARIMAX and GARCH models. Additionally, we extensively discuss temporal aggregation of vector ARMA and multivariate GARCH models. Finally, we address in detail the issue of instantaneous causality spuriously induced by temporal aggregation. Three empirical applications complete the article.

[TABLES 4, 5 AND 6 ABOUT HERE]

In Tables 4, 5 and 6, we provide an overview of the results obtained for the classes of univariate and multivariate models presented throughout the paper. Actually, Tables 4, 5 and 6 include *further* results not surveyed in the previous sections. In particular, Table 4 shows the orders for stock aggregation of several ARIMA models. Table 5 presents the orders of aggregate ARIMAX models, for *all* the flow and

[23]We are grateful to Christian M. Hafner for providing us with the GAUSS code to calculate the temporally aggregated parameters of the MGARCH model in (33) as a function of the aggregation frequency.

stock combinations between the endogenous and the exogenous variables. Table 6 gives the orders for stock aggregation of ARMA-GARCH and vector ARMA models.

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] Bauwens L., Laurent S., Rombouts J.V.K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics* 21, 79-109.

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

[5] Bollerslev T. (1990). Modelling the coherence in short run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics* 72, 498-505.

[6] Bollerslev T., Engle R.F., Wooldridge T. (1988). A capital asset pricing model with time varying covariances. *Journal of Political Economy* 96, 116-131.

[7] Bramati M.C. (2005). Some robust methods in the analysis of multivariate time series. Unpublished Ph.D. dissertation. Université Libre de Bruxelles, Bruxelles, Belgium.

[8] 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.

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

[10] Campbell J.Y., Lo A.W., MacKinlay A.C. (1997). *The Econometrics of Financial Markets.* Princeton: Princeton University Press.

[11] Christiano L.J., Eichenbaum M. (1987). Temporal aggregation and structural inference in macroeconomics. Research Department Working Paper n. 306. Federal Reserve Bank, Minneapolis, MN.

[12] Comte F., Lieberman O. (2000). Second-order noncausality in multivariate GARCH processes. *Journal of Time Series Analysis* 21, 535-557.

[13] Comte F., Lieberman O. (2003). Asymptotic theory for multivariate GARCH processes. *Journal of Multivariate Analysis* 84, 61-84.

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

[15] Drost F.C., Werker B.J.M. (1996). Closing the GARCH gap: Continuous GARCH modeling. *Journal of Econometrics* 74, 31-57.

[16] Engle R.F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroscedasticity models. *Journal of Business & Economic Statistics* 20, 339-350.

[17] Engle R.F., Kroner K.F. (1995). Multivariate simultaneous generalized ARCH. *Econometric Theory* 11, 122-150.

[18] Engle R.F., Ng V.M., Rothschild M. (1990). Asset pricing with a factor ARCH covariance structure: Empirical estimates for treasury bills. *Journal of Econometrics* 45, 213-237.

[19] Florens J.P., Mouchart M. (1985). A linear theory for noncausality. *Econometrica* 53, 157-176.

[20] Forni M., Hallin M., Lippi M., Reichlin L. (2000). The generalised dynamic factor model: Identification and estimation. *The Review of Economics and Statistics* 82, 540-554.

[21] Franses P., van Dijk D. (2000). *Non-linear Time Series Models in Empirical Finance*. Cambridge: Cambridge University Press.

[22] Geweke J.B. (1978). Temporal aggregation in the multiple regression model. *Econometrica* 46, 643-661.

[23] Giacomini R., Granger C.W.J. (2004). Aggregation of space-time processes. *Journal of Econometrics* 118, 7-26.

[24] Gouriéroux C. (1997). *ARCH Models and Financial Applications*. New York: Springer.

[25] Granger C.W.J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica* 37, 424-438.

[26] Granger C.W.J. (1980a). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics* 14, 227-238.

[27] Granger, C.W.J. (1980b). Testing for causality: A personal viewpoint. *Journal of Economic Dynamics and Control* 2, 329-352.

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

[29] Granger C.W.J. (1988). Some recent developments in a concept of causality. *Journal of Econometrics* 39, 199-211.

[30] Granger C.W.J. (1990). Aggregation of time-series variables: A survey. In Barker T. and Pesaran H.M. (eds.) *Disaggregation in Econometric Modelling*. London: Routledge.

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

[32] Granger C.W.J., Morris M.J. (1976). Time series modeling and interpretation. *Journal of the Royal Statistical Society. Series A (General)* 139, 246-257.

[33] Gulasekaran R., Abeysinghe T. (2002). The distortionary effects of temporal aggregation on Granger causality. Working Paper No. 0204. Department of Economics, National University of Singapore, Singapore.

[34] Hafner C.M. (2003). Fourth moment structure of multivariate GARCH models. *Journal of Financial Econometrics* 1, 26-54.

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

[36] Hafner C.M., Herwartz H. (1998). Time-varying market price of risk in the CAPM. Approaches, empirical evidence and implications. *Finance* 19, 93-112.

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

[38] Harvey A.C. (1981). *Time Series Models*. Deddington, Oxford: Allan.

[39] Jeantheau T. (1998). Strong consistency of estimators for multivariate ARCH models. *Econometric Theory* 14, 70-86.

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

[41] Kawakatsu H. (2006). Matrix exponential GARCH. *Journal of Econometrics* 134, 95-128.

[42] Kroner K.F., Ng V.K. (1998). Modeling asymmetric comovements of asset returns. *Review of Financial Studies* 11, 817-844.

[43] Li W.K., Ling S., McAleer M. (2002). Recent theoretical results for time series models with GARCH errors. *Journal of Economic Surveys* 16, 245-269.

[44] Lütkepohl H. (1984). Forecasting contemporaneously aggregated vector ARMA processes. *Journal of Business & Economic Statistics* 2, 201-214.

[45] Lütkepohl H. (1987). *Forecasting Aggregated Vector ARMA Processes*. Berlin: Springer-Verlag.

[46] Lütkepohl H. (2005). *New Introduction to Multiple Time Series Analysis*. Berlin: Springer.

[47] Marcellino M. (1996). Some temporal aggregation issues in empirical analysis. Economics Department working paper 96-39. University of California, San Diego, CA.

[48] Marcellino M. (1999). Some consequences of temporal aggregation in empirical analysis. *Journal of Business & Economic Statistics* 17, 129-136.

[49] Meddahi N., Renault E. (2004). Temporal aggregation of volatility models. *Journal of Econometrics* 119, 355-379.

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

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

[52] Nijman T.E., Sentana E. (1996). Marginalization and contemporaneous aggregation in multivariate GARCH processes. *Journal of Econometrics* 71, 71-87.

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

[54] Pierce D.A., Haugh L.D. (1977). Causality in temporal systems: Characterizations and a survey. *Journal of Econometrics* 5, 265-293.

[55] Proietti T. (2006). On the estimation of nonlinearly aggregated mixed models. *Journal of Computational and Graphical Statistics* 15, 18-38.

[56] Quenouille M.H. (1957). *The Analysis of Multiple Time Series*. London: Griffin.

[57] Renault E., Sekkat K., Szafarz A. (1998). Testing for spurious causality in exchange rates. *Journal of Empirical Finance* 5, 47-66.

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

[59] Silvestrini A., Salto M., Moulin L., Veredas D. (2008). Monitoring and forecasting annual public deficit every month: The case of France. *Empirical Economics*, forthcoming.

[60] Silvestrini A., Veredas D. (2005). Temporal aggregation of univariate time series models. CORE Discussion Paper 2005/59. Université catholique de Louvain, Louvain-la-Neuve, Belgium.

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

[62] Sims C.A. (1972). Money, income and causality. *American Economic Review* 62, 540-562.

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

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

[65] Tiao G.C., Wei W.W.S. (1976). Effect of temporal aggregation on the dynamic relationship of two time series variables. *Biometrika* 63, 513-523.

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

[67] Tse Y.K., Tsui A.K.C. (2002). A multivariate generalized autoregressive conditional heteroscedasticity model with time-varying correlations. *Journal of Business & Economic Statistics* 20, 351-362.

[68] van der Weide R. (2002). GO-GARCH: A multivariate generalized orthogonal GARCH model. *Journal of Applied Econometrics* 17, 549-564.

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

[70] Wei W.W.S. (1982). The effects of systematic sampling and temporal aggregation on causality. A cautionary note. *Journal of the American Statistical Association* 77, 316-319.

[71] Wei W.W.S. (1990). *Time Series Analysis: Univariate and Multivariate Methods*. Redwood City: Addison-Wesley.

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

[73] Zaffaroni P. (2007). Aggregation and memory of models of changing volatility. *Journal of Econometrics* 136, 237-249.

Table 1: Estimated and aggregated models (Belgian Federal Deficit)

	Monthly estimated model	**Quarterly aggregated model**	**Annual aggregated model**
Number of obs.	252	252	252
Model	ARIMA$(0,0,1)(0,1,1)_{12}$	ARIMA$(0,0,1)(0,1,1)_4$	ARIMA$(0,1,2)$
Constant	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
		Quarterly estimated model	**Annual estimated model**
Number of obs.		84	21
Model		ARIMA$(0,0,1)(0,1,1)_4$	ARIMA$(0,1,2)$
Constant		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

Table 2: Univariate GARCH estimation

Coefficient	Estimate	Std. Err.	t-value
$\hat{\psi}$	0.0203	0.0109	1.8624
$\hat{a}$	0.0918	0.0206	4.4563
$\hat{b}$	0.8795	0.0253	34.7628
$\hat{\nu}$	7.8370	1.4886	5.2647
No. Observations	1,303	No. Parameters	4

Univariate GARCH$(1,1)$ model estimated on MSCI Index. The estimated model is $h_t = \hat{\psi} + \hat{a}\varepsilon_t^2 + \hat{b}h_{t-1}$, with the rescaled innovations following a standardized Student distribution with $\hat{\nu}$ degrees of freedom. Estimation sample: from April 1, 2001 to April 30, 2006.

Table 3: Multivariate GARCH estimation

Coefficient	Estimate	Std. Error	t-value
$\hat{\mu}_1$	0.000754	0.00025	3.019656
$\hat{\mu}_2$	0.000659	0.00024	2.748973
$\hat{\hat{C}}(1,1)$	0.001371	0.000136	10.08479
$\hat{\hat{C}}(2,1)$	0.00082	0.000135	6.083377
$\hat{\hat{C}}(2,2)$	0.001269	0.000122	10.36536
$\hat{\hat{B}}(1,1)$	0.965193	0.003946	244.5871
$\hat{\hat{B}}(2,2)$	0.957724	0.004081	234.6605
$\hat{\hat{A}}(1,1)$	0.235672	0.013333	17.67637
$\hat{\hat{A}}(2,2)$	0.259498	0.013338	19.45521
No. Observations	1,693	No. Parameters	9

Multivariate GARCH model estimated on DAX and CAC 40 log returns series. The estimated model is the bivariate BEKK$(1,1,1)$ in (31), with Gaussian innovations. Conditional Mean (same for all series): ARMA$(0,0)$ model. Conditional Variance: Diagonal BEKK$(1,1)$. Estimation Sample: from March 2, 1994 to August 25, 2000.

Table 4: Order Conditions I (ARIMA models)

Model	**Aggr. Model**	**AR**	**S AR**	**MA**	**S MA**
FLOW					
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$(p,d,q) \times (P,D,Q)_s$	ARIMA$(p,d,r) \times (P,D,R)_{s^*}$	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					
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$(p,d,q) \times (P,D,Q)_s$	ARIMA$(p,d,r) \times (P,D,R)_{s^*}$	p	P	$r = \left\lfloor \frac{p(k-1)+d(k-1)+q}{k} \right\rfloor$	$R = \left\lfloor \frac{(P+D)s^*k+(Q-P-D)s}{k} \right\rfloor$

Summary of the orders of the polynomials in the aggregated models. First column is the low frequency model. Second column is the aggregated model. The remaining columns are the others of the aggregated models as a function of the orders of the polynomials of the original models. *Aggr.* stands for Aggregated and S for Seasonal.

Table 5: Order Conditions II (ARIMAX models)

Model	Aggr. Model	AR	MA	Lag Exog.
FLOW+FLOW				
ARIMAX$(p,d,q)(m)$ ARIMA$(v,\tilde{d},w)$	ARIMAX$(p,d,r)(a)$	p	the greater of $\left\lfloor \frac{(p+d+1)(k-1)+q}{k} \right\rfloor$ and $\left\lfloor \frac{(p+d+1)(k-1)+m+\tilde{d}}{k} \right\rfloor + \left\lfloor \frac{(v+\tilde{d}+1)(k-1)+w}{k} \right\rfloor$	$a = \left\lfloor \frac{(p+d+1)(k-1)+m+\tilde{d}}{k} \right\rfloor + v + \tilde{d}$
FLOW+STOCK				
ARIMAX$(p,d,q)(m)$ ARIMA$(v,\tilde{d},w)$	ARIMAX$(p,d,r)(a)$	p	the greater of $\left\lfloor \frac{(p+d+1)(k-1)+q}{k} \right\rfloor$ and $\left\lfloor \frac{(p+d+1)(k-1)+m-1+\tilde{d}}{k} \right\rfloor + \left\lfloor \frac{(v+\tilde{d})(k-1)+w}{k} \right\rfloor$	$a = \left\lfloor \frac{(p+d+1)(k-1)+m-1+\tilde{d}}{k} \right\rfloor + v + \tilde{d}$
STOCK+FLOW				
ARIMAX$(p,d,q)(m)$ ARIMA$(v,\tilde{d},w)$	ARIMAX$(p,d,r)(a)$	p	the greater of $\left\lfloor \frac{(p+d)(k-1)+q}{k} \right\rfloor$ and $\left\lfloor \frac{(p+d)(k-1)+m+\tilde{d}}{k} \right\rfloor + \left\lfloor \frac{(v+\tilde{d}+1)(k-1)+w}{k} \right\rfloor$	$a = \left\lfloor \frac{(p+d)(k-1)+m+\tilde{d}}{k} \right\rfloor + v + \tilde{d}$
STOCK+STOCK				
ARIMAX$(p,d,q)(m)$ ARIMA$(v,\tilde{d},w)$	ARIMAX$(p,d,r)(a)$	p	the greater of $\left\lfloor \frac{(p+d)(k-1)+q}{k} \right\rfloor$ and $\left\lfloor \frac{(p+d)(k-1)+m-1+\tilde{d}}{k} \right\rfloor + \left\lfloor \frac{(v+\tilde{d})(k-1)+w}{k} \right\rfloor$	$a = \left\lfloor \frac{(p+d)(k-1)+m-1+\tilde{d}}{k} \right\rfloor + v + \tilde{d}$

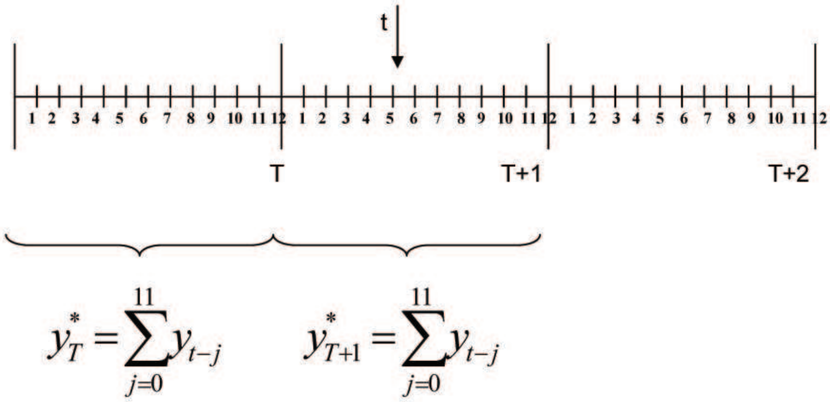
Summary of the orders of the polynomials in the aggregated models with exogenous variables with four different scenarios. FLOW+FLOW means the y_t and x_t are both flow variables. Similarly for the other cases. First column shows the high frequency models for the dependent and exogenous variables. ARIMAX$(p,d,q)(m)$ means that x_t is present in the model through m lags. Second column is the aggregated model for the dependent variable. ARIMAX$(p,d,r)(a)$ means that x_t is present in the model through a lags. The remaining columns are the orders of the aggregated models as a function of the orders of the polynomials of the original models. *Aggr.* stands for Aggregated and *Lag Exog.* for the lag of the exogenous variable in the aggregated model of the dependent variable.

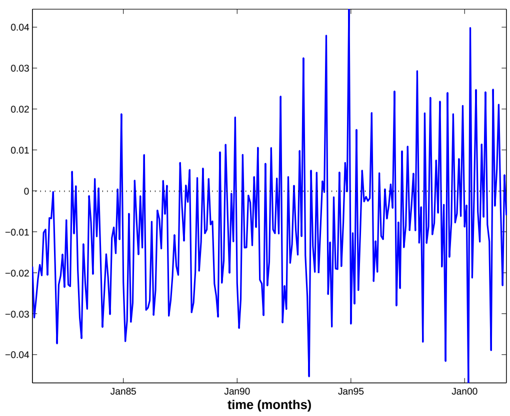
Table 6: Order Conditions III (ARMA-GARCH & VARMA models)

Model	**Aggr. Model**	**AR**	**MA**	
FLOW				
ARMA(p,q)-GARCH(P,Q)	ARMA(p,r)-GARCH(R,R)	p	$r = \left\lfloor \frac{(p+1)(k-1)+q}{k} \right\rfloor$	$R = \tilde{r} + \frac{1}{2}r(r+1)$ $\tilde{r} = \max(P,Q)$
VARMA(p,q)	VARMA(p,r)	p	$r = p - s$ $r = p$ $r = p + 1 + s$	for $sk < p - q + 1 \leq (s+1)k,\ \ s = 0, 1, \ldots, p$ for $p = q$ for $sk < q - 1 - p < (s+1)k,\ \ s = 0, 1, \ldots$
MGARCH$(1,1)$	MGARCH$(1,1)$	1	1	
STOCK				
ARMA(p,q)-GARCH(P,Q)	ARMA(p,r)-GARCH(R,R)	p	$r = \left\lfloor \frac{p(k-1)+q}{k} \right\rfloor$	$R = \tilde{r} + \frac{1}{2}r(r+1)$ $\tilde{r} = \max(P,Q)$
VARMA(p,q)	VARMA(p,r)	p	$r = p - 1 - s$ $r = p$ $r = p + s$	for $sk < p - q \leq (s+1)k,\ \ s = 0, 1, \ldots, p-1$ for $p = q$ for $sk < q - p < (s+1)k,\ \ s = 0, 1, \ldots$

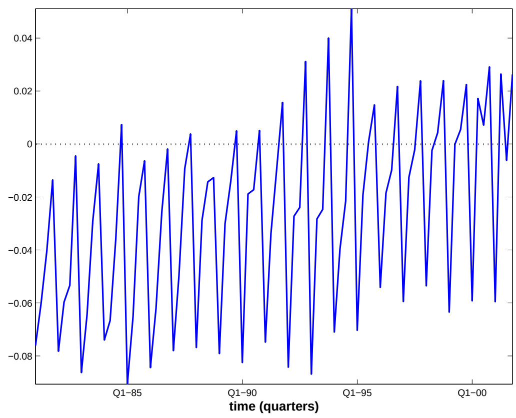
Summary of the orders of the polynomials in the aggregated models. First column is the low frequency model. Second column is the aggregated model. The remaining columns are the others of the aggregated models as a function of the orders of the polynomials of the original models. *Aggr.* stands for Aggregated. Note that, for VARMA models, s is the lowest value such that the inequalities are satisfied.

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

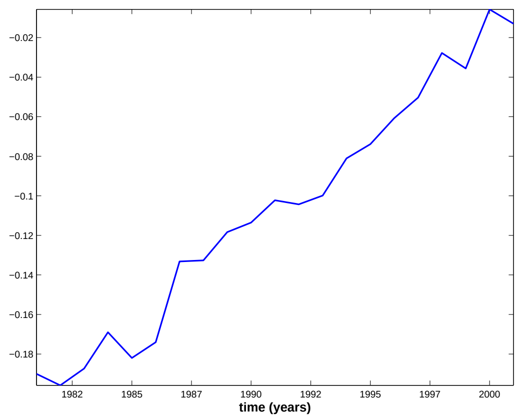




(a)



(b)



(c)

Figure 2: Belgian Federal deficit in real terms at different aggregation frequencies (January 1981-December 2001), y-axis is in Billion Euros: (a) monthly; (b) quarterly; (c) annual.

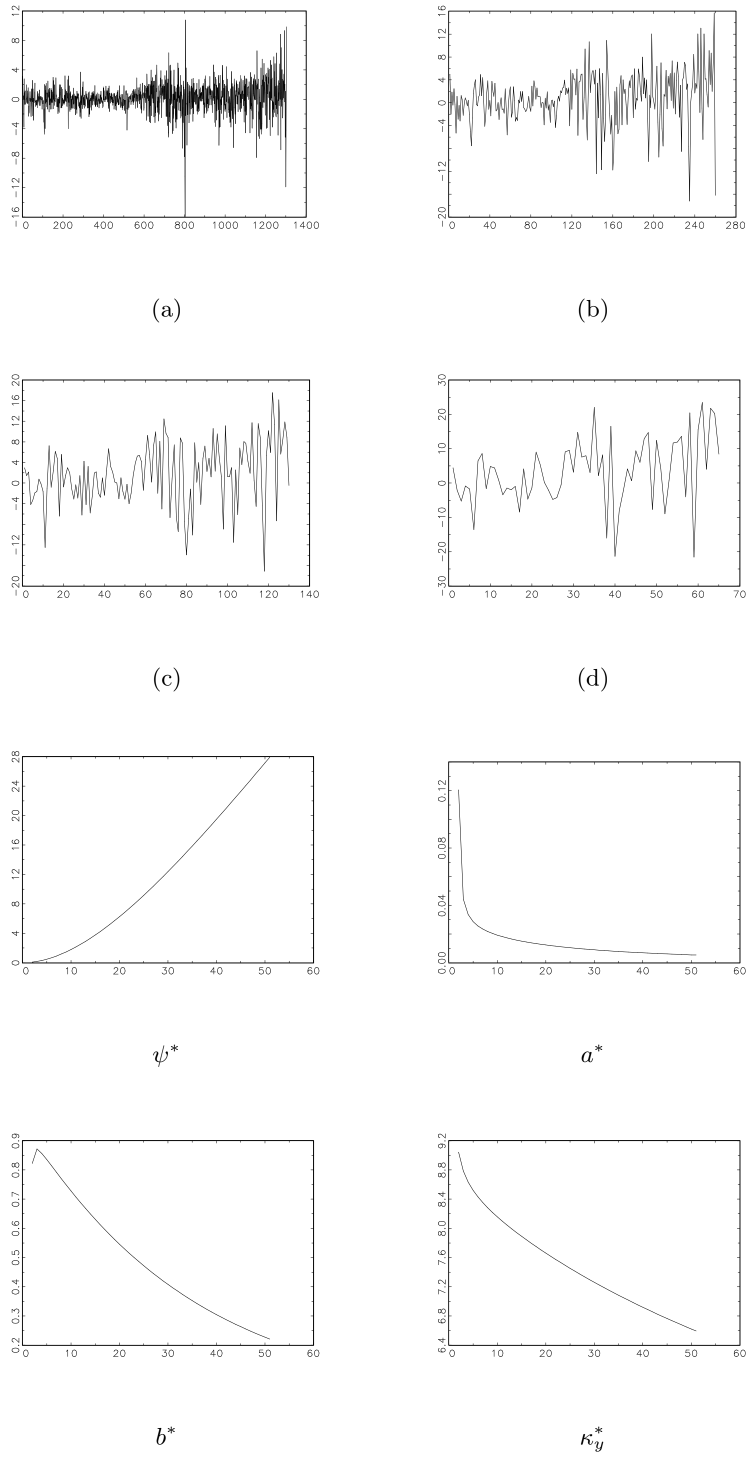
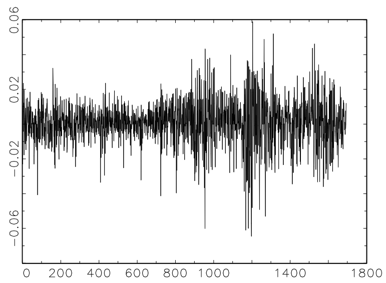
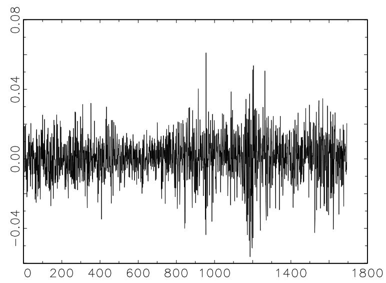


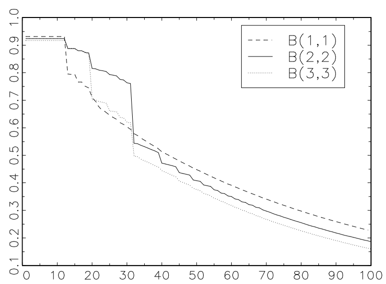
Figure 3: First and second rows show the Morgan Stanley Capital International (MSCI) Index for Indonesia (April 1, 2001 - April 30, 2006), at different aggregation frequencies: (a) daily; (b) weekly; (c) biweekly; (d) monthly. Last two rows show the aggregated GARCH parameters (y-axis) at different frequencies (x-axis): k=2, 3, …, 50.



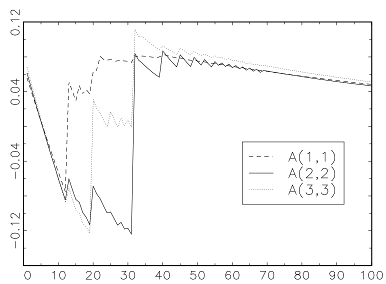
(a)



(b)



(c)



(d)

Figure 4: DAX (a) and CAC 40 (b) log returns series from March 2, 1994 to August 25, 2000. Temporally aggregated elements of the parameter matrices (c) $\mathbf{A}^*$ and (d) $\mathbf{B}^*$ as a function of the aggregation level, for k=2, 3, ..., 100.

RECENTLY PUBLISHED "TEMI" (*)

N. 657 – *Excess money growth and inflation dynamics*, by Barbara Roffia and Andrea Zaghini (January 2008).

N. 658 – *R&D and market structure in a horizontal differentiation framework*, by Davide Fantino (January 2008).

N. 659 – *Housing market spillovers: Evidence from an estimated DSGE model*, by Matteo Iacoviello and Stefano Neri (January 2008).

N. 660 – *Real exchange rate volatility and disconnect: An empirical investigation*, by Riccardo Cristadoro, Andrea Gerali, Stefano Neri and Massimiliano Pisani (April 2008).

N. 661 – *The effect of investment tax credit: Evidence from an atypical programme in Italy*, by Raffaello Bronzini, Guido de Blasio, Guido Pellegrini and Alessandro Scognamiglio (April 2008).

N. 662 – *Accounting for sampling design in the SHIW*, by Ivan Faiella (April 2008).

N. 663 – *Delayed privatization*, by Bernardo Bortolotti and Paolo Pinotti (April 2008).

N. 664 – *Portfolio selection with mononotone mean-variance preferences*, by Fabio Maccheroni, Massimo Marinacci, Aldo Rustichini and Marco Taboga (April 2008).

N. 665 – *Directed matching with endogenous Markov probability: Clients or competitors?*, by Emanuela Ciapanna (April 2008).

N. 666 – *What are borders made of? An analysis of barriers to European banking integration*, by Massimiliano Affinito and Matteo Piazza (April 2008).

N. 667 – *Innovation driven sectoral shocks and aggregate city cycles*, by Andrea R. Lamorgese (April 2008).

N. 668 – *On applying synthetic indices of multidimensional well-being: Health and income inequalities in selected EU countries*, by Andrea Brandolini (April 2008).

N. 669 – *Values, inequality and happiness*, by Claudia Biancotti and Giovanni D'Alessio (April 2008).

N. 670 – *Credit risk and business cycle over different regimes*, by Juri Marcucci and Mario Quagliariello (June 2008).

N. 671 – *Cyclical asymmetry in fiscal variables*, by Fabrizio Balassone, Maura Francese and Stefania Zotteri (June 2008).

N. 672 – *Labour market for teachers: Demographic characteristics and allocative mechanisms*, by Gianna Barbieri, Piero Cipollone and Paolo Sestito (June 2008).

N. 673 – *Output growth volatility and remittances*, by Matteo Bugamelli and Francesco Paternò (June 2008).

N. 674 – *Agglomeration within and between regions: Two econometric based indicators*, by Valter Di Giacinto and Marcello Pagnini (June 2008).

N. 675 – *Service regulation and growth: Evidence from OECD countries*, by Guglielmo Barone and Federico Cingano (June 2008).

N. 676 – *Has globalisation changed the Phillips curve? Firm-level evidence on the effect of activity on prices*, by Eugenio Gaiotti (June 2008).

N. 677 – *Forecasting inflation and tracking monetary policy in the euro area: Does national information help*? by Riccardo Cristadoro, Fabrizio Venditti and Giuseppe Saporito (June 2008).

N. 678 – *Monetary policy effects: New evidence from the Italian flow of funds*, by Riccardo Bonci and Francesco Columba (June 2008).

N. 679 – *Does the expansion of higher education increase the equality of educational opportunities? Evidence from Italy*, by Massimiliano Bratti, Daniele Checchi and Guido de Blasio (June 2008).

N. 680 – *Family succession and firm performance: Evidence from Italian family firms*, by Marco Cucculelli and Giacinto Micucci (June 2008).

N. 681 – *Short-term interest rate futures as monetary policy forecasts*, by Giuseppe Ferrero and Andrea Nobili (June 2008).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2006

F. BUSETTI, *Tests of seasonal integration and cointegration in multivariate unobserved component models*, Journal of Applied Econometrics, Vol. 21, 4, pp. 419-438, **TD No. 476 (June 2003).**

C. BIANCOTTI, *A polarization of inequality? The distribution of national Gini coefficients 1970-1996*, Journal of Economic Inequality, Vol. 4, 1, pp. 1-32, **TD No. 487 (March 2004)**.

L. CANNARI and S. CHIRI, *La bilancia dei pagamenti di parte corrente Nord-Sud (1998-2000)*, in L. Cannari, F. Panetta (a cura di), Il sistema finanziario e il Mezzogiorno: squilibri strutturali e divari finanziari, Bari, Cacucci, **TD No. 490 (March 2004).**

M. BOFONDI and G. GOBBI, *Information barriers to entry into credit markets,* Review of Finance, Vol. 10, 1, pp. 39-67, **TD No. 509 (July 2004).**

W. FUCHS and LIPPI F., *Monetary union with voluntary participation,* Review of Economic Studies, Vol. 73, pp. 437-457 **TD No. 512 (July 2004).**

E. GAIOTTI and A. SECCHI, *Is there a cost channel of monetary transmission? An investigation into the pricing behaviour of 2000 firms*, Journal of Money, Credit and Banking, Vol. 38, 8, pp. 2013-2038 **TD No. 525 (December 2004)**.

A. BRANDOLINI, P. CIPOLLONE and E. VIVIANO, *Does the ILO definition capture all unemployment?*, Journal of the European Economic Association, Vol. 4, 1, pp. 153-179, **TD No. 529 (December 2004)**.

A. BRANDOLINI, L. CANNARI, G. D'ALESSIO and I. FAIELLA, *Household wealth distribution in Italy in the 1990s,* in E. N. Wolff (ed.) International Perspectives on Household Wealth, Cheltenham, Edward Elgar, **TD No. 530 (December 2004)**.

P. DEL GIOVANE and R. SABBATINI, *Perceived and measured inflation after the launch of the Euro: Explaining the gap in Italy,* Giornale degli economisti e annali di economia, Vol. 65, 2 , pp. 155-192, **TD No. 532 (December 2004).**

M. CARUSO, *Monetary policy impulses, local output and the transmission mechanism*, Giornale degli economisti e annali di economia, Vol. 65, 1, pp. 1-30, **TD No. 537 (December 2004).**

L. GUISO and M. PAIELLA, *The role of risk aversion in predicting individual behavior, In P. A. Chiappori e C. Gollier (eds.) Competitive Failures in Insurance Markets: Theory and Policy Implications, Monaco, CESifo,* **TD No. 546 (February 2005).**

G. M. TOMAT, *Prices product differentiation and quality measurement: A comparison between hedonic and matched model methods*, Research in Economics, Vol. 60, 1, pp. 54-68, **TD No. 547 (February 2005)**.

L. GUISO, M. PAIELLA and I. VISCO, *Do capital gains affect consumption? Estimates of wealth effects from Italian household's behavior,* in L. Klein (ed), Long Run Growth and Short Run Stabilization: Essays in Memory of Albert Ando (1929-2002), Cheltenham, Elgar, **TD No. 555 (June 2005).**

F. BUSETTI, S. FABIANI and A. HARVEY, *Convergence of prices and rates of inflation*, Oxford Bulletin of Economics and Statistics, Vol. 68, 1, pp. 863-878, **TD No. 575 (February 2006).**

M. CARUSO, *Stock market fluctuations and money demand in Italy, 1913 - 2003,* Economic Notes, Vol. 35, 1, pp. 1-47, **TD No. 576 (February 2006)**.

S. IRANZO, F. SCHIVARDI and E. TOSETTI, *Skill dispersion and productivity: An analysis with matched data*, CEPR Discussion Paper, 5539, **TD No. 577 (February 2006).**

R. BRONZINI and G. DE BLASIO, *Evaluating the impact of investment incentives: The case of Italy's Law 488/92.* Journal of Urban Economics, Vol. 60, 2, pp. 327-349, **TD No. 582 (March 2006).**

R. BRONZINI and G. DE BLASIO, *Una valutazione degli incentivi pubblici agli investimenti,* Rivista Italiana degli Economisti , Vol. 11, 3, pp. 331-362, **TD No. 582 (March 2006).**

A. DI CESARE, *Do market-based indicators anticipate rating agencies? Evidence for international banks,* Economic Notes, Vol. 35, pp. 121-150, **TD No. 593 (May 2006).**

L. DEDOLA and S. NERI, *What does a technology shock do? A VAR analysis with model-based sign restrictions*, Journal of Monetary Economics, Vol. 54, 2, pp. 512-549, **TD No. 607 (December 2006)**.

R. GOLINELLI and S. MOMIGLIANO, *Real-time determinants of fiscal policies in the euro area*, Journal of Policy Modeling, Vol. 28, 9, pp. 943-964, **TD No. 609 (December 2006).**

2007

S. MAGRI, *Italian households' debt: The participation to the debt market and the size of the loan*, Empirical Economics, v. 33, 3, pp. 401-426, **TD No. 454 (October 2002)**.

L. CASOLARO. and G. GOBBI, *Information technology and productivity changes in the banking industry*, Economic Notes, Vol. 36, 1, pp. 43-76, **TD No. 489 (March 2004)**.

G. FERRERO, *Monetary policy, learning and the speed of convergence,* Journal of Economic Dynamics and Control, v. 31, 9, pp. 3006-3041, **TD No. 499 (June 2004).**

M. PAIELLA, *Does wealth affect consumption? Evidence for Italy,* Journal of Macroeconomics, Vol. 29, 1, pp. 189-205, **TD No. 510 (July 2004).**

F. LIPPI. and S. NERI, *Information variables for monetary policy in a small structural model of the euro area*, Journal of Monetary Economics, Vol. 54, 4, pp. 1256-1270, **TD No. 511 (July 2004)**.

A. ANZUINI and A. LEVY, *Monetary policy shocks in the new EU members: A VAR approach*, Applied Economics, Vol. 39, 9, pp. 1147-1161, **TD No. 514 (July 2004)**.

D. JR. MARCHETTI and F. Nucci, *Pricing behavior and the response of hours to productivity shocks,* Journal of Money Credit and Banking, v. 39, 7, pp. 1587-1611, **TD No. 524 (December 2004).**

R. BRONZINI, *FDI Inflows, agglomeration and host country firms' size: Evidence from Italy*, Regional Studies, Vol. 41, 7, pp. 963-978, **TD No. 526 (December 2004).**

L. MONTEFORTE, *Aggregation bias in macro models: Does it matter for the euro area?*, Economic Modelling, 24, pp. 236-261, **TD No. 534 (December 2004)**.

A. NOBILI, *Assessing the predictive power of financial spreads in the euro area: does parameters instability matter?*, Empirical Economics, Vol. 31, 1, pp. 177-195, **TD No. 544 (February 2005)**.

A. DALMAZZO and G. DE BLASIO, *Production and consumption externalities of human capital: An empirical study for Italy*, Journal of Population Economics, Vol. 20, 2, pp. 359-382, **TD No. 554 (June 2005).**

M. BUGAMELLI and R. TEDESCHI, *Le strategie di prezzo delle imprese esportatrici italiane*, Politica Economica, v. 23, 3, pp. 321-350, **TD No. 563 (November 2005).**

L. GAMBACORTA and S. IANNOTTI, *Are there asymmetries in the response of bank interest rates to monetary shocks?*, Applied Economics, v. 39, 19, pp. 2503-2517, **TD No. 566 (November 2005).**

S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy*, Development Studies Working Papers 231, Centro Studi Luca d'Agliano, **TD No. 570 (January 2006)**.

P. ANGELINI and F. LIPPI, *Did prices really soar after the euro cash changeover? Evidence from ATM withdrawals,* International Journal of Central Banking, Vol. 3, 4, pp. 1-22, **TD No. 581 (March 2006)**.

A. LOCARNO, *Imperfect knowledge, adaptive learning and the bias against activist monetary policies*, International Journal of Central Banking, v. 3, 3, pp. 47-85, **TD No. 590 (May 2006)**.

F. LOTTI and J. MARCUCCI, *Revisiting the empirical evidence on firms' money demand*, Journal of Economics and Business, Vol. 59, 1, pp. 51-73, **TD No. 595 (May 2006)**.

P. CIPOLLONE and A. ROSOLIA, *Social interactions in high school: Lessons from an earthquake*, American Economic Review, Vol. 97, 3, pp. 948-965, **TD No. 596 (September 2006).**

A. BRANDOLINI, *Measurement of income distribution in supranational entities: The case of the European Union*, in S. P. Jenkins e J. Micklewright (eds.), Inequality and Poverty Re-examined, Oxford, Oxford University Press, **TD No. 623 (April 2007).**

M. PAIELLA, *The foregone gains of incomplete portfolios,* Review of Financial Studies, Vol. 20, 5, pp. 1623-1646, **TD No. 625 (April 2007).**

K. BEHRENS, A. R. LAMORGESE, G.I.P. OTTAVIANO and T. TABUCHI, *Changes in transport and non transport costs: local vs. global impacts in a spatial network*, Regional Science and Urban Economics, Vol. 37, 6, pp. 625-648, **TD No. 628 (April 2007).**

G. ASCARI and T. ROPELE, *Optimal monetary policy under low trend inflation*, Journal of Monetary Economics, v. 54, 8, pp. 2568-2583, **TD No. 647 (November 2007).**

R. GIORDANO, S. MOMIGLIANO, S. NERI and R. PEROTTI, *The Effects of Fiscal Policy in Italy: Evidence from a VAR Model*, European Journal of Political Economy, Vol. 23, 3, pp. 707-733, **TD No. 656 (December 2007).**

2008

P. ANGELINI, *Liquidity and announcement effects in the euro area*, Giornale degli Economisti e Annali di Economia, v. 67, 1, pp. 1-20, **TD No. 451 (October 2002).**

F. SCHIVARDI and R. TORRINI, *Identifying the effects of firing restrictions through size-contingent Differences in regulation*, Labour Economics, v. 15, 3, pp. 482-511, **TD No. 504 (June 2004)**.

S. MOMIGLIANO, J. HENRY and P. HERNÁNDEZ DE COS, *The impact of government budget on prices: Evidence from macroeconometric models*, Journal of Policy Modelling, v. 30, 1, pp. 123-143 **TD No. 523 (October 2004).**

P. ANGELINI and A. GENERALE, *On the evolution of firm size distributions*, American Economic Review, v. 98, 1, pp. 426-438, **TD No. 549 (June 2005).**

V. CESTARI, P. DEL GIOVANE and C. ROSSI-ARNAUD, *Memory for Prices and the Euro Cash Changeover: An Analysis for Cinema Prices in Italy*, In P. Del Giovane e R. Sabbatini (eds.), The Euro Inflation and Consumers' Perceptions. Lessons from Italy, Berlin-Heidelberg, Springer, **TD No. 619 (February 2007).**

J. SOUSA and A. ZAGHINI, *Monetary Policy Shocks in the Euro Area and Global Liquidity Spillovers,* International Journal of Finance and Economics, v.13, 3, pp. 205-218, **TD No. 629 (June 2007)**.

M. DEL GATTO, GIANMARCO I. P. OTTAVIANO and M. PAGNINI, *Openness to trade and industry cost dispersion: Evidence from a panel of Italian firms*, Journal of Regional Science, v. 48, 1, pp. 97-129, **TD No. 635 (June 2007)**.

P. DEL GIOVANE, S. FABIANI and R. SABBATINI, *What's behind "inflation perceptions"? A survey-based analysis of Italian consumers,* in P. Del Giovane e R. Sabbatini (eds.), *The Euro Inflation and Consumers' Perceptions. Lessons from Italy*, Berlin-Heidelberg, Springer, **TD No. 655 (January 2008).**

FORTHCOMING

S. SIVIERO and D. TERLIZZESE, *Macroeconomic forecasting: Debunking a few old wives' tales,* Journal of Business Cycle Measurement and Analysis, **TD No. 395 (February 2001)**.

P. ANGELINI, P. DEL GIOVANE, S. SIVIERO and D. TERLIZZESE, *Monetary policy in a monetary union: What role for regional information?*, International Journal of Central Banking, **TD No. 457 (December 2002)**.

L. MONTEFORTE and S. SIVIERO, *The Economic Consequences of Euro Area Modelling Shortcuts,* Applied Economics, **TD No. 458 (December 2002)**.

L. GUISO and M. PAIELLA,, *Risk aversion, wealth and background risk*, Journal of the European Economic Association, **TD No. 483 (September 2003)**.

C. BIANCOTTI, G. D'ALESSIO and A. NERI, *Measurement errors in the Bank of Italy's survey of household income and wealth*, Review of Income and Wealth, **TD No. 520 (October 2004)**.

L. GAMBACORTA, *How do banks set interest rates?,* European Economic Review, **TD No. 542 (February 2005)**.

R. FELICI and M. PAGNINI, *Distance, bank heterogeneity and entry in local banking markets,* The Journal of Industrial Economics, **TD No. 557 (June 2005).**

S. DI ADDARIO and E. PATACCHINI, *Wages and the city. Evidence from Italy,* Labour Economics, **TD No. 570 (January 2006)**.

M. BUGAMELLI and A. ROSOLIA, *Produttività e concorrenza estera*, Rivista di politica economica, **TD No. 578 (February 2006).**

M. PERICOLI and M. TABOGA, *Canonical term-structure models with observable factors and the dynamics of bond risk premia,* **TD No. 580 (February 2006).**

E. VIVIANO, *Entry regulations and labour market outcomes. Evidence from the Italian retail trade sector,* Labour Economics, **TD No. 594 (May 2006)**.

S. FEDERICO and G. A. MINERVA, *Outward FDI and local employment growth in Italy,* Review of World Economics, Journal of Money, Credit and Banking, **TD No. 613 (February 2007).**

F. BUSETTI and A. HARVEY, *Testing for trend*, Econometric Theory **TD No. 614 (February 2007).**

M. BUGAMELLI, *Prezzi delle esportazioni, qualità dei prodotti e caratteristiche di impresa: analisi su un campione di imprese italiane*, Economia e Politica Industriale, **TD No. 634 (June 2007)**.

A. CIARLONE, P. PISELLI and G. TREBESCHI, *Emerging Markets' Spreads and Global Financial Conditions*, Journal of International Financial Markets, Institutions & Money, **TD No. 637 (June 2007).**

S. MAGRI, *The financing of small innovative firms: The Italian case,* Economics of Innovation and New Technology, **TD No. 640 (September 2007)**.

R. BONCI and F. COLUMBA, *Monetary policy effects: New evidence from the Italian flow of funds,* Applied Economics, **TD No. 678 (June 2008)**.

L. ARCIERO, C. BIANCOTTI, L. D'AURIZIO and C. IMPENNA, *Exploring agent-based methods for the analysis of payment systems: A crisis model for StarLogo TNG,* Journal of Artificial Societies and Social Simulation, **TD No. 686 (August 2008)**.