An adaptation of the MIDAS regression model for estimating and forecasting quarterly GDP: Application to the case of Guadeloupe .*

Alain Guay[‡]

Alain Maurin[†]

Cette version: April 2008

Preliminary and incomplete

Abstract

In this article, we propose a new temporal disaggregation technique based on the adaptation of the MIDAS regression for assessing econometric models involving time series data sampled at different frequencies. After an overview of the general features of the MIDAS method, we analyze its relation to the issue of temporal disaggregation. Secondly, we explore a simple disaggregation procedure, which we compare to the more traditional Chow-Lin (1971) and Denton (1971) methods. Subsequently, we discuss the results issued from certain numerical applications realized for the assessment of a quarterly GDP series for Guadeloupe, a French overseas territory for which quarterly accounts have not been compiled and for which there is no available aggregated indicator of economic trends.

*

[‡]Université du Québec à Montréal, CIRPÉE, CIREQ et LEAD. *e-mail*: `guay.alain@uqam.ca`.

[†]Université des Antilles et de la Guyane et LEAD. *e-mail*: `alain.maurin@univ-ag.fr`.

1 Introduction

The need to combine information issued by different publications, whether on a monthly and quarterly basis or on a quarterly and annual basis, was recognized very early on as food for thought by many economic decision-makers and economists. For the former, a group including not only governments and company directors, but also households, the availability of forecast analysis is indispensable for optimum budget management. Taking the most recent information into account during the preparation of analyses and projections, as made possible by the combination of data from different frequencies, is extremely useful. For the latter, the main preoccupation is of a methodological nature. For many years, data generation has focused on the standardization process as the solution to the disaggregation problem. In general terms, the standardization process consists of combining different series of infra-annual and annual reference data, with the aim of obtaining a new infra-annual series. There has been a steady increase in literature specializing in this subject since the early 1960s, with notably a collection of works published by American and other Anglo-Saxon authors, dedicated to the generic notion of benchmarking. Various methods of benchmarking have been developed and widely applied throughout the world in order to generate quarterly data from commonly used economic variables. The Denton (1971) method, which relies principally on optimizing calculations to make the annual totals of infra-annual series tally with annual reference totals, was among the first of its kind and continues to be one of the most widely used benchmarking methods. Using a formalization leading to a solution of generalized least squares, Chow and Lin (1971) also produced a method that has become widely recognized. Today, the development of even more reliable procedures continues to motivate many statisticians and econometrists, as demonstrated by recent publications and studies devoted to this subject (Proietti (2006)). In the domain of forecasting, international literature has also been enriched by diverse contributions proposing a wide range of modelling techniques that associate data from different frequencies in order to calculate projections for the principal economic aggregates. Hence, Shen (1996) used VAR and Bayesian VAR modelling to forecast quarterly variables in Taiwan. By developing a procedure capable of taking into account data issued from monthly series, he demonstrated that this significantly increases forecast accuracy. Chin and Miller (1996), also using the VAR module, proposed another quarterly series forecasting procedure based on the exploitation of monthly data. The latter consists of combining forecasts from two modules estimated using quarterly and monthly data respectively. This article has a twofold aim. On the one hand, it seeks to make a literary contribution by putting forward certain responses to the problem of temporal disaggregation. To this end, we propose a specific econometric procedure based on an adaptation of the MIDAS regression as introduced by Ghysels, Sinko and Valkanov (2006). On the other, we intend to present concrete responses to the need for improved statistical production in developing countries and regions where there is a shortage of temporal data. Indeed, while high-frequency data and studies on economic trends continue to be among the most frequently requested economic analyses in developed countries, they are even more urgently needed in developing countries. In the case of Guadeloupe, there is no need for lengthy justifications when it comes to the importance of empirical reflection and applications dedicated to the study of economic fluctuations (see Maurin and Montauban (2004) and Maurin (2005)). Accordingly, Foy and Maurin (2006) proposed alternative solutions for estimating quarterly GDP series in Guadeloupe. By highlighting that this constitutes a starting condition for the undertaking of studies on economic trends, they launched a debate on the need to establish a reference series for Guadeloupe's quarterly GDP. The methodological discussions and empirical investigations presented in this article are in fact a continuation of the works cited above. More precisely, not only do we intend to explore other alternatives for disaggregating annual GDP data, but we will also examine different methods of forecasting this data. Our work is organized into two parts. The first is dedicated to methodological discussions and reviews the major aspects of the MIDAS method. The second explains the different stages of the empirical applications.

2 Methodology

A regression such as MIDAS allows for the relation between variables sampled at different frequencies to be examined. We shall begin by presenting a simple case. Let us examine the variable y_t sampled once between the period $t-1$ et t. This variable, for instance, can be sampled at a quarterly frequency, the index t would thus correspond to quarters. The explanatory variable x_t^m is sampled m times over the same period. For example, if x_t^m is sampled on a monthly basis, the index m is equal to three. We are seeking here to examine the relationship between y_t and $x_t^{(m)}$. The objective is to project the variable y_t on the historical samplings of $x_{t-j/m}^{(m)}$. The index $t-j/m$ shows that the variable is sampled more frequently than the variable y_t. A MIDAS-type regression is defined by the relation:

$$y_t = \beta_0 + \beta_1 B\left(L^{1/m}; \theta\right) x_t^{(m)} + \varepsilon_t \tag{2.1}$$

where $t = 1, \cdots, T$, $B\left(L^{1/m}; \theta\right) = \sum_{k=0}^{K} B(k; \theta) L^{k/m}$ and $L^{1/m}$ is a lag operator such as $L^{1/m} x_t^{(m)} = x_{t-1/m}^{(m)}$. The term $B\left(L^{1/m}; \theta\right)$ depends on the lag operator $L^{1/m}$ and on a parameter vector of limited dimension θ. The introduction of the vector θ allows for the handling of cases where the number of lags of $x_t^{(m)}$ is considerable. For example, y_t can depend on 12-quarter lags of the explanatory variable $x_t^{(m)}$. Since the latter is sampled on a monthly basis, the number of lags is therefore equal to 36. One approach which would allow us to address the problem of over-parameterization associated with the high number of lags in the case of a regression model with data from different frequencies is to constrain the lag polynomial of the explanatory variable. The latter therefore depends on a limited number of parameters contained here in the vector θ. Ghysels, Sinko and Valkanov (2006) proposed several alternatives for constraining the form of this lag polynomial. In the specification (2.1), the parameter β_1 captures the aggregate effect of lags of $x_t^{(m)}$ on y_t. This parameter can be identified by normalizing the function $B\left(L^{1/m}; \theta\right)$ so that its sum is equal to the unit.

2.1 Polynomial specifications

Specifying the parameters of the polynomial $B\left(L^{1/m}; \theta\right)$ is the major challenge for a MIDAS-type regression model. Ghysels, Sinko et Valkanov (2006) proposed different functional forms for this polynomial. One of these forms is associated with the Almon lag polynomial (Almon, 1965). An exponential variant, it can be defined as follows:

$$B(k; \theta) = \frac{e^{\theta_1 k + \theta_2 k^2 + \ldots + \theta_p k^p}}{\sum_{k=1}^{K} e^{\theta_1 k + \theta_2 k^2 + \ldots + \theta_p k^p}}.$$

This function is very flexible and can take on several configurations, hence its interest. The simple case where $\theta_1 = \theta_2 = \cdots = \theta_p = 0$ corresponds to the case where the lag weight of each lag is the same. This specification corresponds to the often used practice which consists of using an average of monthly values to obtain a quarterly series. In general, the weights can decline more or less rapidly or take on the desired form according to the value of the parameters θ_i, for $i = 1, \cdots, p$.

The second specification proposed by Ghysels, Sinko and Valkanov (2006) depends on two parameters θ_1, θ_2 and has the following form:

$$B(k;\theta_1,\theta_2) = \frac{f\left(\frac{k}{K},\theta_1;\theta_2\right)}{\sum_{k=1}^{K} f\left(\frac{k}{K},\theta_1;\theta_2\right)}$$

and

$$\begin{aligned} f(x,a;b) &= \frac{x^{a-1}(1-x)^{b-1}\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \\ \Gamma(a) &= \int_0^\infty e^{-x}x^{a-1}dx. \end{aligned}$$

This specification is based on the function gamma which is often used in econometrics for its flexibility. The function $f(x,\theta_1;\theta_2)$ can take on several forms, according to the values of θ_1 and θ_2. For example, when $\theta_1 = \theta_2 = 1$, the weights are equal.

The two functions presented above show two important characteristics: (i) the coefficients are positive; (ii) their sum is equal to 1. The fact that the coefficients are positive allows for the volatility process to be assessed and the second characteristic allows for the identification of the parameter β_1. Two issues are yet to be addressed : the choice of K and that of the number of lags to include in the specification. We shall come back to this later on in the article.

Naturally, the specification (2.1) can be generalized. It can be rationalized by including lags of the endogenous variable as well as those of the explicatory variable $x_t^{(m)}$, we would then obtain the following specification:

$$y_t = \mu + \alpha(L)y_{t-1} + \sum_{i=1}^{K}\sum_{j=1}^{L} B_{ij}\left(L^{1/m};\theta\right)x_t^{(m)} + \varepsilon_t \tag{2.2}$$

where $\alpha(L) = \sum_{q=0}^{Q}\alpha_q L^q$. The generalization of the model (2.1) can also be realized in a multivariate context. This will give us:

$$Y_t = \mu + A(L)Y_{t-1} + \sum_{i=1}^{K}\sum_{j=1}^{L} B_{ij}\left(L^{1/m};\theta\right)X_t^{(m)} + \varepsilon_t \tag{2.3}$$

where Y_t, ε_t and X_t are now vectors with μ, $A(L)$ and B_{ij} of compatible dimensions.

2.2 The MIDAS regression and temporal disaggregation

We will show how a disaggregated series can be obtained from a MIDAS-type regression. The approach presented here is therefore an alternative to the more traditional Chow and Lin (1971) and Litterman (1983) approaches. Let us consider the simple case where a y_t series is sampled annually and a $x_t^{(m)}$ series is sampled on a quarterly basis.[1] The relation of the non-sampled quarterly series $y_{t-i/m}^{(m)}$ according to the sampled variable $x_{t-i/m}$ is represented as:

$$y_{t-i/m}^{(m)} = \beta_0^{(m)} + \beta_1^{(m)} \omega_i x_{t-i/m}^{(m)} + \varepsilon_{t-i/m}^{(m)}. \tag{2.4}$$

for $t = 1, \ldots, T$, $i = 0, \ldots, m-1$ and $m = 4$ in this case. It can be noted here that the ω_i weight can vary according to the quarter. In cases where the annual series is the sum of the non-sampled quarterly series, we then have $y_t = \sum_{i=0}^{m-1} y_{t-i/m}^{(m)}$. Using the relation (2.4) we obtain:

$$y_t = \sum_{i=0}^{m-1} y_{t-i/m}^{(m)} = m\beta_0^{(m)} + \beta_1^{(m)} \left(\sum_{i=0}^{m-1} \omega_i x_{t-i/m}^{(m)} \right) + \sum_{i=0}^{m-1} \varepsilon_{t-i/m}^{(m)}.$$

This relation corresponds to a MIDAS regression having the following form

$$y_t = \beta_0 + \beta_1 B\left(L^{1/m};\theta\right) x_t^{(m)} + \varepsilon_t. \tag{2.5}$$

with $\beta_0 = m\beta_0^{(m)}$, $\beta_1 = \beta_1^{(m)}$, $\sum_{i=0}^{m-1} \omega_i x_{t-i/m}^{(m)} = B\left(L^{1/m};\theta\right) x_t^{(m)}$ et $\varepsilon_t = \sum_{i=0}^{m-1} \varepsilon_{t-i/m}^{(m)}$.

In this case, the problem of disaggregation consists of constructing a quarterly series $y_t^{(m)}$ using the annual series y_t and the regression (2.5). Under the hypothesis that the error term is a homoscedastic white noise, a disaggregated series can be obtained using the following transformation:

$$y_{t-i/m}^{(m)} = \frac{1}{m}\hat{\beta}_0 + \hat{\beta}_1 \hat{\omega}_i x_{t-i/m}^{(m)} + \frac{1}{m}\hat{\varepsilon}_t.$$

where $\hat{\beta}_0$, $\hat{\beta}_1$, $\hat{\omega}_i$ are parameter estimators for the MIDAS regression (2.5) and $\hat{\varepsilon}_t$ the residue of this regression. More specifically, the estimators of monthly weights $\hat{\omega}_i$ are provided by the estimator $\hat{B}\left(L^{1/m};\theta\right)$ of the MIDAS regression and the residue $\hat{\varepsilon}_t$ is uniformly distributed for the $i = 0, \ldots, m-1$.

Let us now compare the approach developed here using the MIDAS regression to the Chow-Lin approach. Under the latter, the weighting of the quarterly explicatory variables in the relation (2.4) is assumed to be the same for all quarters, i.e. $\beta_1^{(m)}\omega_i = \beta_1^{(m)}\omega$ for all i imposing an invariant relation according to the quarter. In the case of the MIDAS regression, this weighting is estimated using the polynomial $B\left(L^{1/m};\theta\right)$. The approach proposed here uses a wider range of information than Chow and Lin's more standard method. This additional flexibility should

[1] This presentation can easily be generalized with p explanatory variables.

improve aggregate performance and allow, among other things, for the possible presence of seasonality to be taken into account.

The MIDAS regression also allows for the consideration of autoregressive forms. For instance, the lagged variables of the aggregated series can be included, which can in turn serve to improve the calculation of the disaggregated series. Let us examine MIDAS regression including a lag of the dependent variable. We obtain the following regression equation:

$$y_t = \beta_0 + \rho y_{t-1} + \beta_1 B\left(L^{1/m};\theta\right) x_t^{(m)} + \varepsilon_t. \tag{2.6}$$

This specification can be rewritten in the following form:

$$y_t = \tilde{\beta}_0 + \frac{\beta_1 B\left(L^{1/m};\theta\right)}{(1-\rho L)} x_t^{(m)} + \tilde{\varepsilon}_t. \tag{2.7}$$

We thereby obtain a polynomial having the form $B\left(L^{1/m};\theta\right)\sum_{j=0}^{\infty}\rho^j L^j$. This polynomial is compatible with the seasonal effects of $x_t^{(m)}$ on $y_t^{(m)}$. Indeed, for our example, this polynomial has a moving average representation with periodical effects corresponding to quarters. The MIDAS regression is compatible with a relation between the disaggregated variables as represented by:

$$y_{t-i/m}^{(m)} = \beta_0^{(m)} + \rho y_{t-(m+i)/m}^{(m)} + \beta_1^{(m)} \omega_i x_{t-i/m}^{(m)} + \varepsilon_{t-i/m}^{(m)}. \tag{2.8}$$

We can easily think that it would be more pertinent to examine the relation between the disaggregated series based on a first-order autoregressive representation of the non-sampled disaggregated series. We would then obtain:

$$y_{t-i/m}^{(m)} = \beta_0^{(m)} + \rho^{(m)} y_{t-(1+i)/m}^{(m)} + \beta_1 \omega_i x_{t-i/m}^{(m)} + \varepsilon_{t-i/m}^{(m)}. \tag{2.9}$$

where the variable $y_{t-i/m}^{(m)}$ depends on its non-sampled value from the preceding quarter. This specification can then be rewritten in the following form:

$$y_{t-i/m}^{(m)} = \tilde{\beta}_0^{(m)} + \frac{\beta_1 \omega_i}{(1-\rho L^{1/m})} x_{t-i/m}^{(m)} + \tilde{\varepsilon}_{t-i/m}^{(m)}. \tag{2.10}$$

By aggregating this relation, we obtain:

$$y_t = \sum_{i=0}^{m-1} y_{t-i/m}^{(m)} = m\tilde{\beta}_0^{(m)} + \frac{\beta_1^{(m)}\left(\sum_{i=0}^{m-1}\omega_i x_{t-i/m}^{(m)}\right)}{(1-\rho L^{1/m})} + \sum_{i=0}^{m-1}\tilde{\varepsilon}_{t-i/m}^{(m)}.$$

This relation is consistent with a MIDAS regression having the form

$$y_t = \beta_0 + \frac{\beta_1 B\left(L^{1/m};\theta\right) x_t^{(m)}}{(1-\rho L^{1/m})} + \varepsilon_t. \tag{2.11}$$

We thereby obtain a polynomial depending on $L^{1/m}$ having the form $B\left(L^{1/m};\theta\right)\sum_{j=0}^{\infty}\rho^j L^{j/m}$. This specification represents a challenge in terms of estimation since the lagged variable is not sampled. We can however use a finite approximation of the polynomial ratio in (2.11) and an instrumental variable estimation method when the number of samples is substantial.

Chow and Lin's (1971) approach also allows us to take the autocorrelation into account by correcting the efficient least squares estimator in order to obtain a generalized least squares estimator. This correction is achieved using a first order autoregressive process for the error term of the disaggregated series. Unlike the approach proposed here, this temporal dependence does not serve to improve the calculation of the disaggregated series, but simply to obtain an effective estimator of β_0 and β_1. Accordingly, the disaggregated series obtained by Chow and Lin (1971) for the representation (2.5) is given in matrix form by:

$$\hat{Y}^{(m)} = X^m\hat{\beta} + V\mathcal{C}\left(\mathcal{C}'V\mathcal{C}\right)^{-1}\mathcal{C}\hat{\varepsilon}_t^{(m)}.$$

where X^m is a matrix containing a constant and the disaggregated samples x_t^m, V is the variance-covariance matrix that takes into account the presence of autocorrelation and heteroscedasticity in the disaggregated error terms and $\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1)'$. The matrix $\mathcal{C}$ with a dimension of $T \times Tm$ has the following written form:

$$\mathcal{C} = \begin{bmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & . & . & . & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & . & . & . & 0 & 0 & 0 & 0 \\ . & . & . & . & . & . & . & . & . & . & . & . & . & . & . \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & . & . & . & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Thus, the estimators of $\beta = (\beta_0, \beta_1)'$ are the generalized least squares estimators defined as follows:

$$\hat{\beta} = \left[X'(\mathcal{C}V\mathcal{C}')^{-1}X\right]^{-1}X'(\mathcal{C}V\mathcal{C})^{-1}Y$$

with $X = \mathcal{C}X^m$, $Y = \mathcal{C}Y^{(m)}$ and $\hat{\varepsilon} = \mathcal{C}\hat{\varepsilon}^{(m)} = Y - X\hat{\beta}$. Based on the Chow and Lin approach, Litterman (1983) proposed a correction which would take autocorrelation into account in a more complex manner by introducing a random operation boosted by an AR(1) process for the disaggregated error term.

3 Application to the case of Guadeloupe

3.1 Stakes surrounding the quarterisation of annual GDP

Today, there are many databases for aggregates and other indicators outlining the evolution of Guadeloupe's economic activity. Yet, they remain affected by a global shortage in terms of both quantity and quality. Indeed, a rapid overview of the available variables shows that there is an abundance of quarterly, monthly and annual data, collected by various organisms such as the INSEE (French national statistics office), the IEDOM (representative of the French Central Bank in overseas departments), the Chamber of Commerce and Industry, the customs division, the tax authorities, the industrial, research and environmental authorities, the port authority, etc. These organisms fulfil a certain number of missions, including the diffusion of statistical information to companies, local authorities and to the wider public in general. It is therefore possible to track the evolution of economic and monetary activities in quite a few sectors. In spite of this, a close examination of these databases reveals that there is a shortage of information, which leads to certain consequences. First of all, with respect to annual data, the absence of economic aggregates expressed in volume as well as that of certain key variables represents a definite barrier for any attempt at econometric modelling for Guadeloupe (see Maurin and Montauban (2004)). Then, with regards to the short-term data that interests us in this article, the situation seems even more critical - there is no available quarterly or monthly synthetic indicator of economic activity. Maurin (2005) discussed the state of affairs of the currently available economic time series as well as the urgent need to improve the organization of collection methods.

In terms of quality, close examination also highlights the limits of statistical capacities when faced with the need for short-term information among various decision-makers and stakeholders. In most developed economies, quarterly accounts serve as a source of information and as an indispensable tool for economic analysis, since they constitute the sole set of coherent indicators available in the short term and capable of providing a global indication of recent economic activity, in both financial and non-financial domains (European Commission, 1995). In Guadeloupe, quarterly accounts are for the most part absent, as is an index of industrial production. Yet, there is no shortage of internationally adopted daily practices to illustrate the necessity of short-term economic data. For example, after the economic and financial crises which shook Asia in 1997 and Argentina in 2001, the chaotic economic situations and risks of widespread bankruptcy demonstrated to what extent economic fluctuations, and their recessional phases in particular, can lead to extensive material and social difficulties in developing countries. Moreover, among the official responses adopted in the wake of these economic crises, governments were obliged to invest in the development of systems for generating economic data. According to an article published on the United Nations website

The economic crisis in Asia in 1997 has created a demand for short-term economic indicators among other economic and financial data to monitor more closely the economy. Quarterly accounts have become one of the focuses of the program of statistical development and improvement in the ASEAN region, which includes ten member countries: Brunei, Cambodia, Indonesia, Laos, Malaysia, Myanmar, The Philippines, Singapore, Thailand and Vietnam. Many workshops and training sessions on annual and quarterly accounts and related statistics have been carried out by the Asian Development Bank, IMF and by the UNSD within the framework of the UN Statistical Development Account.

Generally speaking, the use of this data can be classed into two domains:

o the regulation of economic activity: cyclical, quarterly and monthly data represent the basic components of preparatory studies for determining economic policy. These studies allow public authorities to optimize their fund allocation in order to better regulate economic activity.

o the answer to the needs of private sector players: compared to governmental decision-makers, private sector players are extremely diverse and comprise central banks, commercial bank research departments, financial markets, potential investors, households, etc. Studies on economic trends are a crucial factor in their decisions to hire, invest, save, etc.

3.2 Data

AnnualGDP: two annual series are available for Guadeloupe. The first covers the period 1970-1994, in base 1990, and contains data expressed in millions of francs. The second, drafted using the new 1995 base, assembles data for the period 1993-2005 expressed in euros. Data for the sub-period 1993-2002 is definitive and reliable, as indicated by the French national statistical institute l'INSEE on its internet site: "To generate new accounts for the French Departments of America (FDA), estimation methods have been updated and new sources have been exploited. Furthermore, modernized national accounting concepts and a new classification system for activities and products, consistent with the corresponding European classification, have been implemented. These major conceptual, methodological and statistical changes have lead to significant revisions of DFA accounts and of their principal aggregates."

Quarterly indicators: since the aim of this interpolation exercise is to constitute a data series that would provide the most realistic overview of the wealth produced each quarter within the economy, it is imperative that we select variables which can potentially serve as standards. Logically, this selection must be based on relevant economic criteria as well as on the availability of statistics. We have therefore reviewed the statistical series generated by different administrative bodies in Guadeloupe and examined how the economic and statistical elements of each series can best reflect the ups and downs of economic activity. We have drawn up a list of twenty-odd potential variables, including production, tax and external trade variables as well as monetary and financial indicators. We have also taken account of several features relating to the behaviour of certain variables that are generally recognized as lag, coincident or lead indicators of economic activity. For example, Nilsson and Brunet (2005), who rated 148 indicators for six major non-OECD economies (Brazil, China, India, Indonesia, Russia and South Africa), showed that:

o the balance of opinion series are generally lead variables;
o real-sector indicators (production, construction, transport, labour, prices) in 50
o external trade variables are often coincident;
o financial variables are rarely lead but often coincident.

In light of these elements, we have retained the following two indicators:

- VAT revenues. Value added tax represents the biggest source of indirect taxes. By definition, it is a general consumption tax applicable to all good and services delivered in France. Since VAT implicates all economic players within the scope of their business operations, the sum of these revenues is an accurate reflection of the level of economic activity. Today, it is widely accepted that VAT revenues react to current cyclical trends whereas direct taxes, which are calculated based on the preceding year, tend to show a lagged reaction. Indeed, in 2004 the French budget minister highlighted that "regardless of the scholarly debates among economic analysts, VAT returns are the best indicator of economic health." Economists within the ministry have also noted that in statistical terms, the quarterly variables of tax revenues show a certain level of seasonality, which, in essence, can be explained by their collection rate. Accordingly, with regard to the State's budgetary situation, on August 31 2006, the French ministry of economy and finance published the following statement on its website:

"The seasonality of tax revenues is largely due to the recovery rate of direct taxes. Excluding cases of monthly deduction, income tax is collected in February, May and September. Corporate tax is collected in four instalments (March, June, September and December), with a balance payment in April. Indirect taxes show more regularity, although VAT revenues are traditionally higher in certain months (January, April, August, and October)."

In Guadeloupe, VAT is collected by both the general treasury office and the customs department, on behalf of the State, and VAT revenues constitute one of the indicators monitored in the IEDOM quarterly bulletin.

- Octroi de mer (dock duties)

  This is an import duty levied in the French overseas departments by the customs department on behalf of the State. The sums collected are re-distributed to regional and municipal authorities. For regional authorities, the octroi de mer is a veritable instrument of economic policy. By controlling the variations of the different rates imposed on each product, it gives itself a certain amount of leeway for influencing the behaviour of economic agents such as household consumption and corporate importation. In late 2004 for example, Guadeloupe's Regional Council removed the octroi de mer from infant formula, milk and baby food. On the other hand, within the framework of a new health policy, it increased duties on certain products, notably on alcohols. In more general terms, the regional authority has declared that it is in favour of using the octroi de mer as a lever for finding a balance between the promotion of local production, the protection of Guadeloupean consumers and the improvement of financial resources within its municipalities.

  Statistics on this variable are available (from 1993 onward) and are regularly published in the IEDOM quarterly bulletin.

3.3 Applications and empirical results

Our empirical investigations were carried out using two software programmes designed for numerical and econometric calculations. Using the MATLAB package, we constituted a database comprising various functions and programmes that use different stages of the adaptation of the MIDAS regression to calculate disaggregated data. We also used a RATS programme to apply the conventional Chow-Lin procedure and two of its derived methods (Fernandez (1981) and Litterman (1983)). Charts 1-3 below are graphic representations of annual GDP figures and quarterly data on VAT and OM duties. The grouping of these unadjusted series may give rise to a few doubts concerning the harmonization of quarterly data, which is not very smooth, with the linear trajectory of annual GDP. The issue of dealing with seasonal effects upstream of the application of disaggregation procedures therefore arises here. There are two arguments in favour of resorting to the corrected data of seasonal variations. First, it is widely believed that traditional methods, such as the Chow-Lin procedure, are not particularly suitable for applications to seasonal data. Secondly, even if non-corrected data is used, it follows that the disaggregated series obtained will inherit the seasonal features of the chosen quarterly indicator. This a priori judgement on the seasonal aspects of quarterly GDP is unsubstantiated. On the contrary, we have highlighted that the MIDAS procedure is remarkably capable of taking the possible presence of seasonal data movements into account. We shall, for these reasons and according to the individual cases, retain the corrected data of seasonal variations or of unadjusted data for VAT and OM indicators.

Furthermore, authors such as Di Fonzo (2003) and Mazzi and Savio (2005) have discussed the shortcomings of conventional procedures based on regression equations, highlighting among other things that the β vector of parameters obtained in the Chow-Lin method is inconsistent when the residues are generated by a regression carried out on non stationary series. In light of these observations on disaggregation data and methods, we can proceed with the generation of our numerical applications.

Based on our two indicators, we have tested several configurations of the regression equation of the Chow-Lin model as well as relation (2.6) of the MIDAS method. A formal rule must be implemented in order to carefully select a final solution among the temporal series generated by these regressions. This in mind, we have chosen to compare the values of the likelihood function associated with each model. The highest value indicates the model that has best adjusted the data sampled. Table 1 lists the results obtained using conventional methods. For models using VAT as the sole explicatory variable, or OM, or both, it should be noted that the method that achieved the best performances was the Litterman approach. Based on the hypothesis of non-stationarity of the residual term of the Chow-Lin model, the latter proposes an AR(2) specification with a unit root.

3.4 Conclusion

Literature devoted to data development and dissemination has been arousing renewed interest since the late 1990s. Indeed, in the wake of exceptional economic and institutional events, such as the financial crises in Asia and Latin America, the adoption of the euro and the strengthening of European construction, analysts and policymakers have naturally been brought to express a stronger demand for data and studies on business cycles. Temporal disaggregation techniques were at the time proposed as solutions to improve the availability of monthly and quarterly estimates of key economic variables. With the contributions of pioneer statisticians such as Chow and Lin and Denton in 1971 and the recent contributions of econometricians like Proietti in 2006, today, there are a few dozen disaggregation algorithms with varying complexity and performance levels. The new procedure that we propose in this article can be included among these methods. Compared to the latter, it presents certain originalities and strong points. It is based on the MIDAS regression, which is intrinsically a modelling approach that directly links variables sampled at different frequencies. As a result, it can calculate estimators using a richer set of information than that used in conventional disaggregation methods. Furthermore, based on a mathematical framework that provides interesting statistical properties, such as the treatment of seasonality in initial data, our procedure should generate better results than the standard methods identified in literature.

References

[1] Almon, S. (1965), “ The distributed Lag between Capital Appropriations and Expenditures,” *Econometrica,* 33, 178-196.

[2] Chin D.M., Miller P.J., 1996, Using monthly data to improve quarterly model forecasts, Federal Reserve Bank of Minneapolis Review, vol. 20, No 2, Spring 1996, pp. 16-33.

[3] Chow G.C., Lin An-Loh (1971), Best linear unbiased interpolation, distribution and extrapolation of times series by related series, Review of Economic and Statistics, Vol. 53, p. 372-375.

[4] Denton F.T. (1971), Adjustment of monthly or quarterly series to annual totals: an approach based on quadratic minimization, Journal of the American Statistical Association, Vol. 66, p. 92-102.

[5] Di Fonzo T. (2003), Temporal disaggregation using related series: log-transformation and dynamics extensions, RISEC, Volume 50 (2003), No. 3, 371-400. Fernandez, R.B. (1981), A methodological note on the estimation of time series,” The Review of Economics and Statistics 63(3), pp. 471-476.

[6] Foy Y., Maurin A., 2006, En quête dun indicateur dactivité trimestrielle pour la Guadeloupe. Partie I: La trimestrialisation du PIB: méthodes et résultats, Documents de travail, LEAD

[7] Ghysels, E., A. Sinko and R. Valkanov (2006), ”MIDAS Regressions: Further Results and New Directions”, Econometric Review, forthcoming.

[8] Ingenito R., Trehan B., 1996, Using monthly data to predict quarterly output, Federal Reserve Bank of San Francisco Economic Review, No 3, 1996, pp. 1-11.

[9] Litterman, R.B. (1983), ”A random walk, markov model for the distribution of time series,” Journal of Business and Statistics, 1, pp. 169-173.

[10] Maurin A. (2005), Quel dispositif dindicateurs pour lanalyse conjoncturelle en Guadeloupe? Document de travail, LEAD.

[11] Maurin A., Montauban J-G. (2004), "La modélisation macroéconomique appliquée au cas des ROM : bilan et discussion sur les problèmes théoriques et méthodologiques", in Maurin A., Montauban J-G. et Vellas F. (éds), Lenjeu du développement économique insulaire : les régions ultra-périphériques de lunion européenne et les TOM, Préface de Brigitte Girardin, Ministre de lOutremer, éditions SEDES, Ouvrage collectif, Juillet 2004.

[12] Mazzi G., Savio G. (2005), Theory and applications of univariate and multivariate models for temporal disaggregation, Mimeo, Statistical Office of the European Communities, Eurostat.

[13] Nilsson R. and Brunet O., 2005, Composite Leading indicators for major OECD non-member economies : Brasil, China, India, Indonesia, Russian Federation, South Africa, OECD Statistics Working Paper, December.

[14] Proietti T., 2006, Temporal disaggregation by state space methods: dynamic regression methods revisited, Econometrics Journal, vol. 9, pp. 357-372

[15] Shen C. H., 1996, Forecasting macroeconomic variables using data of different periodicities, International journal of forecasting, No 12, pp. 269-282.

Figure 1: PIB: annual series

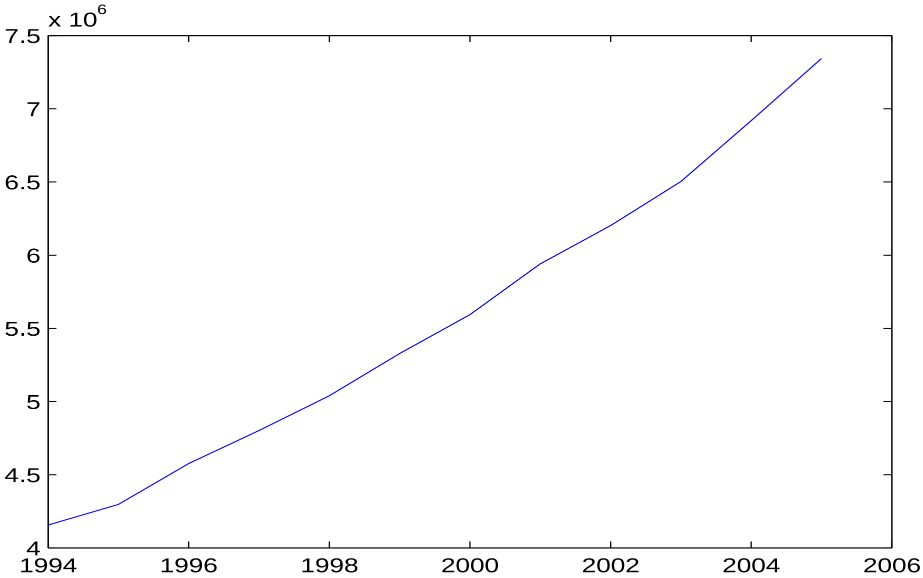

Figure 2: VAT: quarterly series

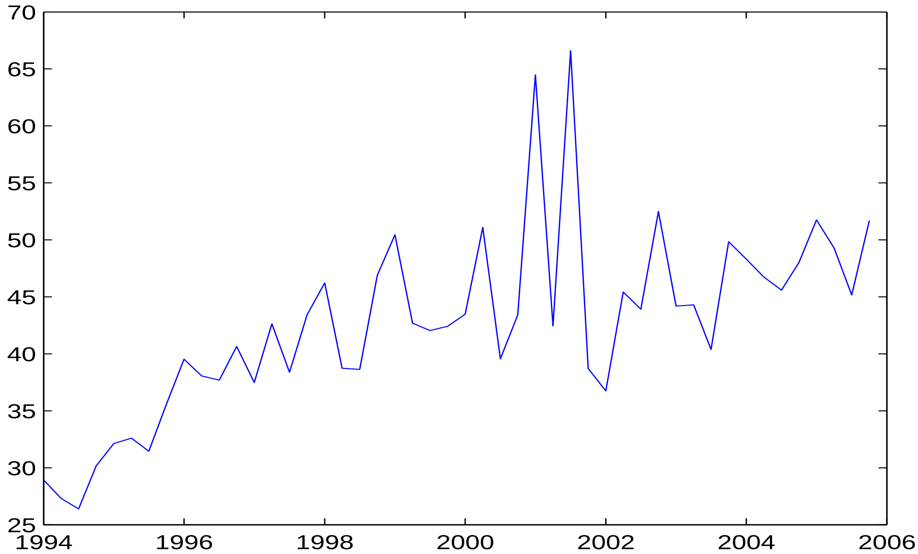

Figure 3: Dock Duties: quarterly series

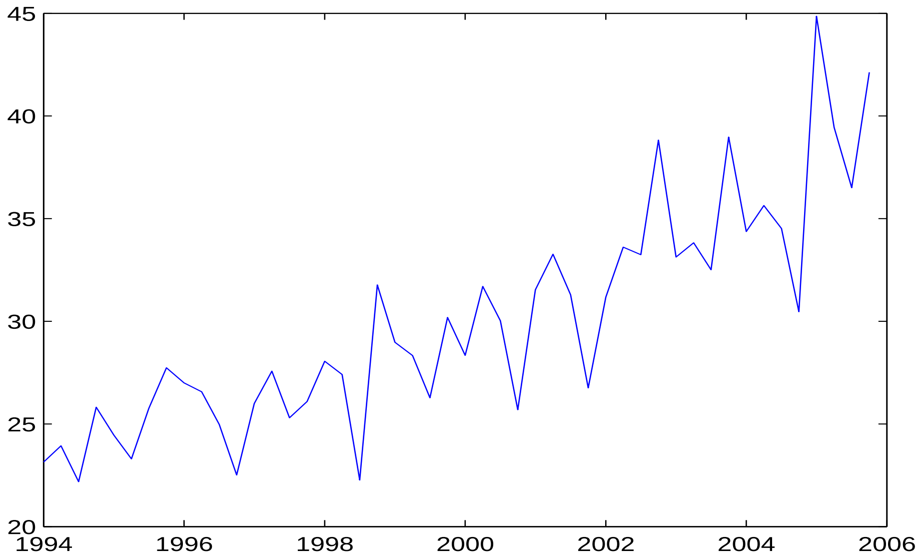

Figure 4: Disaggregated series, AR(1) specification