

تقدير الناتج المحلي الإجمالي الفلسطيني باستخدام نماذج الانحدار للبيانات ذات الترددات المختلطة

*Forecasting Palestinian Gross Domestic Product Using Mixed Data Sampling Regression Techniques*

Student Name and Number:

Mustafa M. Al-Qawasmi / 1115280

Discussion Date : June, 7 , 2014

Committee:

Dr Tariq Sadiq

Supervisor

Dr Abd Al Hakeem Eideh

Dr Wael Al-Hashlamoon

*This Thesis was Submitted in Partial Fulfillment of the Requirements for the Master's Degree in Applied Statistics from the Faculty of Graduate Studies at Birzeit University. Palestine*



تقدير الناتج المحلي الإجمالي الفلسطيني باستخدام نماذج الانحدار للبيانات ذات الترددات المختلطة

*Forecasting Palestinian Gross Domestic Product Using Mixed Data Sampling Regression Techniques*

Student Name and Number:

Mustafa M. Al-Qawasmi / 1115280

Discussion Date : June, 7 , 2014

Committee:

Name	Signature
Dr Abd Al Hakeem Eideh	_____
Dr Tariq Sadiq	_____
Dr Wael Al-Hashlamoon	_____

## *Dedication*

*To those who have supported me at all the time.....*

*To my mother, To my Father,*

*To my wife,*

*To my son and daughter,*

*To my Brothers and Sisters,*

*To my Family,*

*To my colleagues.*

*With all my love and respect*

## *Acknowledgments*

I would like to express my deepest gratitude to my supervisor, Dr. Tariq Sadiq who suggested the subject of my thesis and who made great effort to guide me throughout my studies. This project would not have been possible without his continuous guidance and support. I would like to express my deep appreciation and gratitude to both Dr. Abd Al Hakeem Eideh and Dr. Wael Al-Hashlamoon for their contributions and encouragement. I gratefully acknowledgment the help of Mr. Eric Ghysels who have encouraged me to study mixed data sampling regressions and supported me his data and working files. I would like to thank and express my appreciation to my colleagues in work and study. I would like to express my appreciation to all doctors taught me during my study. Last but not least, my deep thanks go to my family (specially my wife) for their endless encouragement and support and for everyone helped me or encouraged me.

## **Table of Contents**

Table of Contents .....	V
Abbreviations .....	VII
Abstract .....	VIII
ملخص.....	IX
Chapter 1 : Introduction.....	1
1.1 Background .....	1
1.2 Literature Reviews .....	3
1.3 Problem And Methodology .....	5
Chapter 2 : MIDAS Regressions.....	7
2.1 The Forecasting Environment And Notations .....	7
2.2 Data Form .....	8
2.3 Time Aggregation.....	9
2.3.1 Time-Averaging.....	10
2.3.2 Step Weighting(U-MIDAS).....	11
2.3.3 MIDAS Regressions.....	12
2.4 Simple R-MIDAS Regression .....	12
2.5 Polynomial Specifications .....	14
2.5.1 Exponential Almon Lag Polynomial.....	14
2.5.2 Beta Lag function.....	16
2.5.3 Almon Lag Polynomial of order (P) .....	16
2.6 Lag Length Selection Criteria .....	18
2.7 Testing Unit Root and Cointegration .....	19
2.7.1 Unit Root .....	19
2.7.2 Cointegration.....	21

2.7.3 Johansen's procedure .....	22
2.8 Cointegration-MIDAS and ECM-MIDAS models .....	25
2.9 Distributed-Lag Models .....	26
2.10 Estimation Methods .....	30
2.10.1 The Method of Least squares.....	30
2.10.2 The Method of Maximum Likelihood.....	32
2.10.3 Generalized Method Of Moments Estimation.....	34
2.10.4 The Method of Nonlinear Least squares(Gauss-Newton method).....	35
2.11 Forecasting.....	38
Chapter 3 : Application Case Study .....	40
3.1 Modeling The Palestinian GDP .....	40
3.2 Discussion of Results and Conclusions .....	47
Appendix A : Output Tables .....	49
Appendix B : Time Series Plots.....	56
Appendix C : Testing The Flat Aggregation Scheme(AGK test) .....	58
Appendix D : The R-Codes Used In Analysis.....	59
References.....	65
Index.....	70

## **Abbreviations**

**MIDAS: Mixed Data Sampling.**

**U-MIDAS: Unrestricted Mixed Data Sampling.**

**R-MIDAS: Restricted: Mixed Data Sampling.**

**ADL: Autoregressive Distributed Lag.**

**ARIMA: Autoregressive Integrated Moving Averages.**

**GDP: Gross Domestic Product.**

**CIM: Cost of Imports.**

**CPI: Consumer Price Index.**

**EII: Employment In Israel.**

**EALP: Exponential Almon Lag Polynomial.**

**GMM: General Method of Moments.**

**NLS: Non-linear Least Squares.**

**MLE: Maximum-Likelihood Estimator.**

**MSE: Mean Square Error.**

**AIC: Aikaike's Information Criterion.**

**BIC: Bayesian Information Criterion.**

**ADF: Augmented Dickey-Fuller.**

**ECM: Error Correction Model.**

**VECM: Vector Error Correction Models.**

**OLS: Ordinary Least Squares.**

**VAR: vector Autoregression.**

**QMLE: Quasi Maximum Likelihood Estimator.**

**PMA: Palestine Monetary Authority.**

**EIS: Employment rate in Israel and Settlements.**

**ARDL: Auto Regressive Distributed Lag.**

**PCBS: Palestinian Central Bureau of Statistics.**

## **Abstract**

In the case of using classical linear regression models for time series, researchers usually deal with equal frequencies for all the variables. They cannot directly apply such models to a mixed-frequency dataset. The Mixed Data Sampling (MIDAS) regression models deal with this type of data; typically the economic indicators from those observed daily, monthly, quarterly to those yearly. In this study, we introduce MIDAS regression approach which is relatively assumed as a new area. We will explain its ability of dealing with mixed frequency data, and its efficiency of improving parameters estimation and forecasting performance in the presence of extreme observations. To the best of author's knowledge, this is the first research that examines the relationship between the real GDP in Palestine and other indicators using MIDAS regressions. The classical temporal aggregation method is compared to the two types of MIDAS regressions; the restricted and the unrestricted, to build both long-run and short-run relationships. The study results exhibited that both U-MIDAS and R-MIDAS were better than the classical Time-Averaging method in reducing forecasting errors.

The results exhibited that the quarterly Palestinian GDP in the long-run affected by its cost of imports in the second month computed in each quarter, also, the quarterly GDP has increasing general trend and affected by its first lag and the quarterly Employment Rate In Israel and Settlements, and all these variables have significant positive relationship with GDP. In the short-run, the results showed that the quarterly Palestinian GDP is affecting by the second-month and the third-month of the cost of imports of Palestine.



## ملخص

في حالة استخدام نماذج الانحدار الخطي الكلاسيكي في دراسة العلاقات بين السلاسل الزمنية فإن الباحثين عادة ما يتعاملون مع ترددات متساوية لجميع المتغيرات، في حين أنه لا يمكن تطبيق مثل هذه النماذج مباشرة في حالة وجود مجموعة بيانات ومتغيرات تردداتها مختلطة وغير متساوية. اعتمدت هذه الرسالة على دراسة نوع خاص من نماذج الانحدار التي يمكنها أن تتعامل مع مثل هذا النوع من البيانات و السلاسل الزمنية ذات الترددات المختلطة مثل المؤشرات الاقتصادية التي يتم معاينة بعضها يوميا والبعض الآخر يتم معاينتها شهريا أو ربعيا أو سنويا. تعرض هذه الدراسة نماذج انحدار البيانات ذات التكرارات المختلطة (نماذج ميداس) والتي تعتبر من المواضيع الحديثة نسبيا. من خلال هذه الدراسة سوف يتم شرح قدرة هذه النماذج على التعامل مع البيانات ذات الترددات المختلطة، حيث سيتم شرح خصائصها وكفاءتها في تحسين التقدير والتنبؤ للمتغيرات الاقتصادية، كما سيتم مقارنة هذه النماذج بنوعها (المقيد وغير المقيد) مع أسلوب التجميع الزمني الكلاسيكي حسب الأوساط وذلك من خلال بناء النماذج التي تمثل العلاقات بين المتغيرات الاقتصادية على المدى الطويل والعلاقات على المدى القصير، حيث أظهرت نتائج الدراسة تفوق نماذج ميداس غير المقيدة والمقيدة في تقليل أخطاء التنبؤ على نماذج التجميع الزمني الكلاسيكي حسب الأوساط .

حسب معرفة المؤلف فإن هذا هو البحث الأول الذي يدرس العلاقة بين الناتج المحلي الإجمالي الحقيقي في فلسطين وغيرها من المؤشرات الاقتصادية باستخدام نماذج انحدار البيانات ذات التكرارات المختلطة ميداس، حيث أظهرت نتائج الدراسة وجود علاقة دالة إحصائيا على المدى الطويل بين الناتج المحلي الإجمالي الربعي في فلسطين وكل من تكلفة الواردات المحسوبة في ثاني شهر من كل ربع، ومعدل العمالة في إسرائيل والمستوطنات، كما أظهرت النتائج وجود دالة متزايدة للناتج المحلي الربعي في فلسطين على المدى الطويل وتأثره ذاتيا بقيمه السابقة أو المتأخرة. من ناحية أخرى أظهرت نتائج الدراسة وجود علاقة دالة إحصائيا على المدى القصير بين الناتج المحلي الإجمالي الربعي في فلسطين و تكلفة الواردات المحسوبة في أول و ثاني شهر من كل ربع .

## **Chapter 1 : Introduction**

### **1.1 Background**

The analysis of time series data is a pivotal issue. The most of time series are economic or financial variables that affect and are affected by each other. Studying such variables is with a great deal of importance. Both economic and financial variables give us information about the economic level that was reached by any country. Studying such variables help to investigate the effects caused by some indicators on the economic activity in general, then to assist the decision making process.

In order to study the relationships between these variables, the most efficient widely used method is to build regression models. Regression models of time series variables can be univariate or multivariate models. The univariate models investigate the relationship between the variable and itself; in other words, with its time delays(lags) like Autoregressive Integrated Moving Averages(ARIMA) models. From the other hand, the multivariate models investigate the relationship between dependent (response) variable and independent variables(indicators or predictors). Multivariate regression models may investigate the effect of response lags with other predictors on the response, so in this case the model is called autoregressive model.

In building model stage, it is supposed that data(dependents or independents) are frequency-unified or have fixed frequency; this means that each observation of the dependent variable is matched with an observation of the independent variable/s; in other words, each observation of the dependent corresponding to an observation of the independent/s. In this case, it is simply possible to build regression models to investigate the relationship between these variables. Let's suppose that the gross domestic product(GDP) and the cost of imports(CIM) in Palestine are computed monthly by the Palestinian Central Bureau of Statistics; in this case each observation of GDP corresponds to an observation of CIM; the GDP in January is matched with the CIM in January; also, the GDP in February is matched with the CIM in February, and so on for the rest of months and years. So we can easily construct regression model to investigate the relationship between the GDP as dependent (response) and the CIM as independent(predictor).

In fact, most of research and statistics centers in the world cannot provide monthly information about GDP, but they can for CIM. They can just provide quarterly information about GDP, so in this case we have economic variables with different frequencies(mixed frequencies). The GDP has quarterly frequency

observations, but the CIM has monthly frequency observations. It is clear that the frequency of the CIM observations is larger than the frequency of the GDP observations. We call the GDP as low-frequency variable and the CIM as high-frequency variable. We can imagine such mixed frequency data(mixed-data-sampled) as a mixture of observed and missing data; each monthly observation of CIM does not have corresponding observation of GDP since the GDP is quarterly sampled. So the form of data can be imagined by thinking of availability of CIM observations as three times per quarter. In the first quarter of any year, the observation of CIM in January has no corresponding observation of GDP in January; also the observation of CIM in February has no corresponding observation of GDP in February, but the observation of CIM in March has corresponding observation of GDP in March. So we have two missing observations of GDP in both January and February but not in March, whereas there are no missing observations of CIM in the three months. In the same way we have missing observations of GDP in both April and May but not in June, whereas there are no missing observations of CIM in the three months in the second quarter, and so on for the rest of months and quarters of the year.

Sometimes, economic and financial data and variables are available only on an annual basis. For example if we assume that the GDP is sampled annually in the end of each year, and the CIM is sampled monthly; so in this case, we have eleven missing observations of GDP from January to November; while the CIM observations are available completely for the twelve months in that year. So the question that arises here is how can we investigate the relationship between variables that are sampled with mixed frequencies like the GDP as dependent and the CIM as independent?.

The availability of data sampled at different frequencies always presents a problem for researchers working with time series data. If one deleted all cases containing the missing observations, then he would lose much important information, because variables that are available at high frequency contain potentially valuable information[1,4]. The simplest idea that comes into mind after that is to aggregate the high-frequency observations by taking the average of them, and then, the high-frequency variable becomes as the low-frequency variable. For instance, taking the average of the three monthly observations of the CIM in each quarter to get one averaged observation matched with the quarterly observation of the GDP; but it is clear, by this way that we assign the same importance level for the partitions of the high-frequency variable. Therefore, we lose important monthly information, and the results of the analysis will be less accurate[18].

The previous explanations can be summarized by aggregating the high-frequency observations like the monthly observations of the CIM series, and converting it into a quarterly observations to be matched with the quarterly GDP data. Instead of this method, the common solution in such cases is to estimate missing observations of the low-frequency variable like the quarterly GDP. This can be done by using the Kalman filter process, and these subjects are outside the framework of this thesis[4,8,10,18]. Nowadays, Some scientists have proposed methods which can deal with variables sampled at different frequencies such as monthly and quarterly data. Such methods have many advantages and properties comparing with old and common methods had known. We call these methods Mixed Data Sampling(MIDAS) Regressions.

## **1.2 Literature Reviews**

We Always find statisticians interested in getting good properties and improvement on estimations that can be yielded by traditional high level methods such as ARIMA and regression models that deal with time series data. Ghysels, Santa-Clara and Valkanov in 2004 have recently proposed regression models that directly accommodate variables that are sampled at different frequencies. These regression models combining for example monthly and quarterly data together, or quarterly with annually data. Their Mixed Data Sampling – or MIDAS – regressions represent simple, parsimonious, and flexible class of time series models. They showed that their MIDAS regressions will always lead to more efficient(less error-variance) estimation than the classical approach that aggregate all series to be in the same frequencies.

Ghysels, Santa-Clara and Valkanov[1] used many functional forms dealing with few number of parameters in order to estimate MIDAS models. They examined the asymptotic properties of MIDAS regression estimation, and they discussed challenging econometric issues. Clements and Galvao[5, 22] investigated whether the MIDAS approach of Ghysels can be successfully adapted to the short term forecasting of output growth, given that it has been used for forecasting financial variables with daily observations. They extended the distributed lag MIDAS specification to include an autoregressive term (the MIDAS-AR). Sinko[2] discussed two lag parameterizations of MIDAS regressions; the Exponential Almon and Beta. He introduced several new MIDAS specifications that include more general mixed-data structures of unequally spaced observations. Klaus[4] in his thesis compared the forecasting success of two new mixed-frequency time series models: the mixed-frequency VAR and the MIDAS approach. During his work, he outlined the data

generating processes (DGP) that have been used for model comparisons. Klaus listed some conclusions about model specification of mixed-frequency time series models. Due to him, forecast performance, lag length selection and restrictions of the weighting functions in the MIDAS framework are interrelated. He concluded that whether the weighting functions should be restricted or not, depends on the data structure.

Andreou, Ghysels and Kourtellis[8] covered MIDAS regressions and the relationship with the Kalman filter. They discussed DL-MIDAS(Distributed Lag Mixed Data Sampling) models with autoregressive high-frequency predictors. Armesto, Engemann and Owyang[18] compared between Time- Aggregation techniques which include Time-Averaging, Step Weighting and MIDAS regressions. They also offered some forecasting issues such as the End-Of-Period forecasting and the Intra-Period forecasting. Miguel[20] introduced robust procedure that pre-cleans the data by exponential smoothing before the MIDAS regression constructed. His results showed that in the presence of outliers, utilizing the robust version of exponential smoothing to clean the contaminated dataset prior to use MIDAS reduces forecast errors consistently. Chen and Tsay[9] presented a generalized autoregressive distributed lag (GADL) model in order to conduct regression estimations that involve mixed-frequency data. They offer comparisons between the relative performance of the OLS-based GADL specifications and the NLS-MIDAS(Nonlinear Least Squares Mixed Data Sampling models). Forni, Marcellino and Schumacher[6] discussed the unrestricted MIDAS(U-MIDAS) models based on a simple linear lag polynomial. They showed that U-MIDAS work better than MIDAS method if the differences in sampling frequencies are small.

Goetz, Hecq and Urbain[21] proposed in their paper, mixed frequency error correction model in order to model non-stationary variables that are possibly co-integrated. Forni, Marcillino & Shumacher[3] studied the performance of MIDAS models which does not depend on the functional distributed lag polynomials. They compared U-MIDAS models with MIDAS models, and they showed that U-MIDAS performs better than MIDAS in the case of small differences in sampling frequencies. Forni & Marcillino[3] showed by simulation experiments and actual data, that the use of mixed-frequency data, combined with a proper estimation approach, can alleviate the temporal aggregation bias, mitigate the identification issues, and yield more reliable policy conclusions. Forni and Marcellino[7] clarified that MIDAS models appear to be more robust than bridge equation models and the state-space approaches in miss-specification, and computationally simpler.

On the empirical side, Shi[17] studied the effect of lag numbers on MIDAS regressions. He concluded that there exists seasonal effect of the lag numbers in different horizons. He noted that the out-of-sample and in-sample results of realized volatility(RV) and realized absolute volatility(RAV) are quite similar, but in sometimes, out-of sample performs better. Shi stated several findings from predictability of daily to monthly realized volatility of Chinese market. Asimakopulos, Paredes and Warmedinger[19] concentrated on forecasting annual budgetary executions and their subcomponents issues. They assessed the effect of intra-annual fiscal data on the annual outturn of disaggregated series. They employed MIDAS approach to analyze mixed frequency fiscal data. They concluded that timely and good quality data are powerful tools to improve forecasts.

### **1.3 Problem And Methodology**

In this thesis, we will study the problem of specifying the factors affect the Palestinian Gross Domestic Product(GDP) which is sampled quarterly, where some of these probable factors are sampled monthly and the others are sampled quarterly, and we will forecast the future quarterly real GDP using these mixed sampled data using MIDAS regression approach without resorting to fix data frequencies or dealing with missing data.

The MIDAS approach may relate the observations of the low-frequency variable to the lagged observations of the high-frequency variable by distributed lag functions. The proper choice of the functional form, such as the exponential distributed lag polynomial, allows assigning few number of parameters to large number of lags in order to obtain a parsimonious model which implies to simply interpretation of results[18]. As an alternative to such parameterization, we will discuss an unrestricted MIDAS(U-MIDAS), based on a simple linear lag polynomial. This specification has the advantage of a higher flexibility compared to the functional lag polynomials in the standard MIDAS approach[18]. However, U-MIDAS has disadvantage that it need to estimate a lot of parameters if the lag order is large, this leads to difficulty of interpretations. On the other hand, U-MIDAS can be helpful when differences in sampling frequencies are small[6]. For example, monthly-quarterly data; such as if quarterly GDP should be related to monthly cost of imports, and generally a small number of lags is necessary to capture the dynamic relations.

We will explain some models, and other methods will be discussed in the study; such as the usage of MIDAS regressions in building Cointegrating relationships, and the equilibrium (error) correction models, in order to study long-run and short-run

relationships between variables. We will improve the discussion of all topics by applications and examples supported by data tables, reports and summaries that can expand our understanding of the subject.

In next sections, we discuss simple time-averaging, step-weighting function(U-MIDAS), and the restricted MIDAS regressions(R-MIDAS). We will compare between these methods by modeling Palestinian GDP on some important factors using the three methods. We will use MIDAS regressions to forecast the future value of quarterly Palestinian GDP. The main model in this study is defined as:

$$LGDP_t^{(Q)} = \beta_0 + \beta_1 t + \beta_2 LGDP_{t-1}^{(Q)} + \beta_3 LEII_t^{(Q)} + \sum_{k=0}^{m-1} \beta_{k,4} LCPI_{t-\frac{k}{m}}^{(m)} + \sum_{i=0}^{m-1} \beta_{i,5} LCIM_{t-\frac{i}{m}}^{(m)} + u_t$$

Where:

$LGDP_t^{(Q)}$ : the log of the Quarterly Gross Domestic Product as low-frequency Variable.

t: time as low-frequency Variable.

$LGDP_{t-1}^{(Q)}$ : the log of the first lag of the Quarterly Gross Domestic Product as low-frequency Variable.

$LEII_t^{(Q)}$ : the log of the Quarterly Employment in Israel and Settlements as low-frequency Variable.

$LCPI_{t-\frac{k}{m}}^{(m)}$ : the log of the Monthly Consumer Price Index as high-frequency Variable.

$LCIM_{t-\frac{i}{m}}^{(m)}$ : the log of the Monthly Cost of Imports as high-frequency Variable.

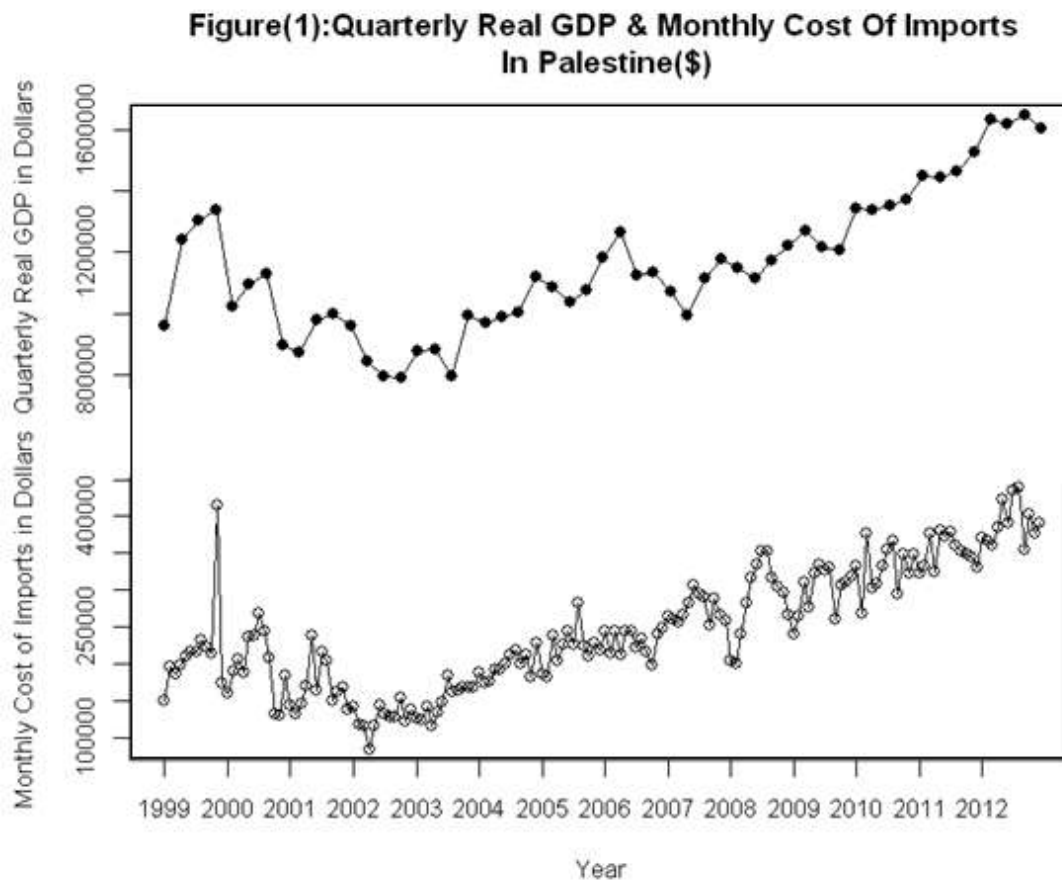
$u_t$ : is the white Noise Process(independent identically distributed random variables with zero mean and constant variance).

Finally, we will use the same frequencies of the high-frequency Variables that are monthly, we will not use weekly or daily high-frequency Variables.

## **Chapter 2 : MIDAS Regressions**

### **2.1 The Forecasting Environment And Notations**

The problem of mixed sampling frequencies is exemplified in Figure (1) which shows quarterly real GDP and monthly cost of imports for the period 1999 to 2012\*. From the figure, we note that monthly cost of imports observations fluctuate between quarterly GDP observations (monthly observations are more concentrated through quarterly observations into the same time period)[18].



When comparing across modeling environments, it is important to use common notations. In the econometric procedures that we will follow, our objective is to forecast a lower-frequency variable ( $y$ ); sampled at periods denoted by time index  $t$ . Past realizations of the lower-frequency variable are denoted by the lag operator,  $L$ .

---

\* The data used in plotting figure(1) were collected from Palestinian Central Bureau of Statistics(PCBS).



For example, if  $y$  is the quarterly GDP, then the GDP one quarter prior would be the first lag of  $y_t$ ,  $Ly_t = y_{t-1}$ , two quarters prior would be  $L^2y_t = y_{t-2}$ , and so on. In addition, we are interested in the information contents of the higher-frequency variable  $x$  which is sampled  $m$  times between the sample periods of  $y$  (between  $t-1$  and  $t$ ), so we will denote to  $x$  as  $x_t^{(m)}$ .  $L^{\frac{k}{m}}$  denotes the lag operator for the higher-frequency variable. If  $x_t^{(m)}$  is the monthly cost of imports used to forecast quarterly GDP denoted by  $y_t$ , then  $L^{\frac{k}{m}}x_t^{(m)} = x_{t-\frac{k}{m}}^{(m)}$  denotes the cost of imports of the  $k$ 'th previous month in the same quarter[18].

Figure(2): Forecast Time Line

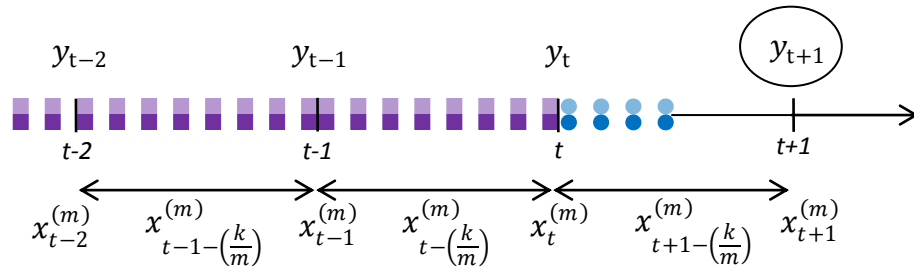


Figure (2) represents the forecast timeline, which for simplicity shows one-period-ahead forecasts. Assume that at time  $t$  we are interested in forecasting  $y_{t+1}$  (the circled observation on the timeline). Standard forecasting experiments would use data available up to time  $t$ ; this is depicted in the dashed section of the timeline. The dotted section of the timeline depicts information that becomes available during the  $t+1$  period; which is called leads, these information may be relevant for forecasting  $y_{t+1}$ . Using MIDAS methods, we can perform intra-period forecasting experiments using both data specified in the dashed section and that in the dotted section[18].

## 2.2 Data Form

Before modeling high and low frequency data, it is important to know how is the form of these variables. Table (1) illustrates the mixed data scheme by the help of the quarterly-monthly example. In the quarterly-monthly mixed data, we have ( $m=3$ ), that is we have three observations of the high-frequency variable between the period  $t-1$  and  $t$ . For two years, we have ( $n=24$ ) monthly high-frequency observations of  $x$  (12 in each year), and we have ( $T=8$ ) quarterly low-frequency observations of  $y$  (4 in each year), so we have  $n=m*T=3*8=24$  known observations of  $x$  and only 8 known observations of  $y$ . If we assumed that the first observation of  $x$  represents the value of

CIM(Cost of Imports) in January/1999  $\left(x_{1-\frac{2}{3}}^{(3)}\right)$ , then the third observation of  $x$ , which represents the value of CIM in March/1999  $(x_1^{(3)})$ , will correspond to the first observation of  $y$  which represents the value of GDP in the first quarter  $(y_1)$ , and so on. For monthly-daily data, we have  $(m=22)$  assuming fixed number of days per month, and we can generalize the same data scheme or form for monthly-daily, annually-quarterly or any other form. The following table shows quarterly-monthly data structure.

**Table (1): Quarterly-Monthly Data Structure for two years(1999-2000)**

$t$	Year/Quarter	First month $x_{t-\frac{2}{m}}^{(m)}$	Second month $x_{t-\frac{1}{m}}^{(m)}$	Third month $x_t^{(m)}$	$y_t$
1	1999/(first quarter)	$x_{1-\frac{2}{3}}^{(3)}$	$x_{1-\frac{1}{3}}^{(3)}$	$x_1^{(3)} = x_1$	$y_1$
2	1999/(second quarter)	$x_{2-\frac{2}{3}}^{(3)}$	$x_{2-\frac{1}{3}}^{(3)}$	$x_2^{(3)} = x_2$	$y_2$
3	1999/(third quarter)	$x_{3-\frac{2}{3}}^{(3)}$	$x_{3-\frac{1}{3}}^{(3)}$	$x_3^{(3)} = x_3$	$y_3$
4	1999/(fourth quarter)	$x_{4-\frac{2}{3}}^{(3)}$	$x_{4-\frac{1}{3}}^{(3)}$	$x_4^{(3)} = x_4$	$y_4$
5	2000/(fifth quarter)	$x_{5-\frac{2}{3}}^{(3)}$	$x_{5-\frac{1}{3}}^{(3)}$	$x_5^{(3)} = x_5$	$y_5$
6	2000/(sixth quarter)	$x_{6-\frac{2}{3}}^{(3)}$	$x_{6-\frac{1}{3}}^{(3)}$	$x_6^{(3)} = x_6$	$y_6$
7	2000/(seventh quarter)	$x_{7-\frac{2}{3}}^{(3)}$	$x_{7-\frac{1}{3}}^{(3)}$	$x_7^{(3)} = x_7$	$y_7$
8	2000/(eighth quarter)	$x_{8-\frac{2}{3}}^{(3)}$	$x_{8-\frac{1}{3}}^{(3)}$	$x_8^{(3)} = x_8$	$y_8$

### **2.3 Time Aggregation**

In this section we will describe three methods of Time-Aggregation of the higher-frequency data that can be used in forecasting lower-frequency variables, that are: Time-Averaging, Unrestricted MIDAS and Restricted MIDAS.

### 2.3.1 Time-Averaging

The first and the simplest method of Time-Aggregation is Time-Averaging[18]. The method of converting higher-frequency data to match the observations of the lower-frequency data by computing the simple average of the observations of the high-frequency variable  $x$  that occur between samples of the lower-frequency variable  $y$  such that:

$$\bar{x}_t = \frac{1}{m} \sum_{k=0}^{m-1} L^{\frac{k}{m}} x_t^{(m)} ; \text{ where } L^{\frac{k}{m}} x_t^{(m)} = x_{t-\frac{k}{m}}^{(m)}$$

Here, we are assign each lag of  $x$  by the same weight or coefficient in each quarter which is  $(1/m)$ . For example by table (1) we have :

$$\bar{x}_8 = \frac{1}{3} \sum_{k=0}^{3-1} L^{\frac{k}{3}} x_t^{(3)} = \frac{1}{3} (x_8^{(3)} + x_{8-\frac{1}{3}}^{(3)} + x_{8-\frac{2}{3}}^{(3)})$$

Letting  $k=0$  and for quarterly-monthly data(i.e.  $m=3$ ), we have  $L^{\frac{k}{m}} x_t^{(m)} = x_{t-\frac{k}{m}}^{(m)} = x_t^{(3)}$ , which means the last-month observation of  $(x)$  in the quarter  $t$ . For  $k=1$  we have  $L^{\frac{k}{m}} x_t^{(m)} = x_{t-\frac{1}{3}}^{(3)}$ , which means the previous month of the last-month(i.e. the second month) observation of  $(x)$  in the quarter  $t$ , and for  $k=2$  we have  $L^{\frac{k}{m}} x_t^{(m)} = x_{t-\frac{2}{3}}^{(3)}$ , which means the first-month observation of  $(x)$  in the quarter  $t$ .

The regression model of  $y_t$  on its own lags with  $\bar{x}_t$  and its lags( $p$  lags for  $y$  and  $q$  lags of  $\bar{x}$ ) becomes clearly identified as:

$$y_t = \alpha + \sum_{i=1}^p \beta_i L^i y_t + \sum_{j=0}^q \gamma_j L^j \bar{x}_t + u_t \quad (1)$$

Where  $\gamma_j$  are the coefficients of the time-averaged  $x$ 's, and  $u_t$  is the white noise process (independent identically distributed random variables with zero mean and constant variance) . The model in equation (1) called autoregressive distributed lag model of order  $p$  and  $q$  (ADL( $p,q$ ))(see section 2.9)[8,51].

The method of time-averaging is very simple and easy to conduct since it is depending on time series regression approaches, but it is clear that this method make a lot of potentially useful information discarded because of equal-weights assigning, thus rendering the relation between the variables will be less accurate[2,18]. Finally,

it is more appropriate to take averages for stock variables, but for flow variables, the higher-frequency values are simply added[4].

### **2.3.2 Step Weighting(U-MIDAS)**

The other method of time aggregation is to give each lag of  $x$  different weight or coefficient in each quarter[18]. We call this method step-weighting or unrestricted mixed data sampling regression(U-MIDAS). The autoregressive distributed lag U-MIDAS(ADL( $p,m$ )-UMIDAS) model can be written as:

$$y_t = \alpha + \sum_{i=1}^p \beta_i L^i y_t + \sum_{k=0}^m \gamma_k L^{\frac{k}{m}} x_t^{(m)} + u_t \quad (2)$$

Where  $u_t$  is the white noise process as defined before. For the purposes of forecasting, the previous equation can be expanded for quarterly-monthly variables when  $m=3$  as:

$$y_{t+1} = \alpha + \beta_1 y_t + \beta_2 y_{t-1} + \dots + \beta_p y_{t-p-1} + \gamma_1 x_t^{(m)} + \gamma_2 x_{t-\frac{1}{3}}^{(m)} + \gamma_3 x_{t-\frac{2}{3}}^{(m)} + u_t \quad (3)$$

Equation(3) represents forecasting equation for the period of  $t+1$  of  $y$ . It is clear that this equation does not contain leads, but if it should, we can add lead terms, and those are:  $\gamma_1^* x_{t+1}^{(m)} + \gamma_2^* x_{t-\frac{1}{3}}^{(m)} + \gamma_3^* x_{t-\frac{2}{3}}^{(m)}$ . This means that we have high-frequency observations now of the variable ( $x_t$ ) in the period ( $t+1$ ). For example, in our research, we analyzed quarterly-monthly data from 1999 to 2012(56 quarterly observations of ( $y$ ) and 168 monthly observations of ( $x$ )). We can forecast  $y$  at the first quarter in 2013 directly by equation(3) without leads. But if we now in the third month of the first quarter in 2013; and the observations of  $x$  at January, February and March are available now; then we can insert the three lead terms in the equation to get benefits of leads information(see section 2.11 for more details about forecasting)[6,18].

It is clear from the equation (3) that despite step-weighting model preserves the timing information(because the high-frequency observations between the periods  $t$  and  $t-1$  were taken into account and not just averaged), but it violates the parsimony principle. In the case of quarterly-monthly variables, it seems that there is no violation of the parsimony principle; because we need only three coefficients to estimate for the variable  $x$ . But for monthly-daily variables it is different; we will need 22 coefficients for the variable  $x$  alone, this is called a parameter proliferation

problem. Also, Model (3) can be extended to multiple lags, then the number of parameters will become quite large[6,18].

### **2.3.3 MIDAS Regressions**

The model which could solve the problem of parameter proliferation while preserving some timing information is called MIDAS model or the restricted mixed data sampling regression(R-MIDAS)[18]. This model was constructed by Ghysels, Santa-Clara, and Valkanov[1]. This model can be written as:

$$y_t = \alpha + \sum_{i=1}^p \beta_i L^i y_t + \gamma \sum_{k=0}^m B(k, \boldsymbol{\theta}) L^{\frac{k}{m}} x_t^{(m)} + u_t \quad (4)$$

Where  $B(k, \boldsymbol{\theta})$  is a polynomial function that determines the weights for temporal aggregation, and the vector  $\boldsymbol{\theta}$  have a specific number of parameters  $\theta_i$ ,  $u_t$  is white noise process[8,18].

Equation(4) represents an autoregressive distributed lag model(ADL-RMIDAS) because  $y_t$  is regressed on its own lags with  $x_t$  and its lags as mentioned before. Returning to equation (2), we note that it is the same as equation (4). Only the difference between them is that the weighting terms of the higher-frequency variable  $x_t$  are  $\{\gamma_j\}$  in equation (2) while they are  $\{B(j, \boldsymbol{\theta})\}$  in equation (4). The important property of the weighting polynomial function  $B(j, \boldsymbol{\theta})$  is that it contains limited number of parameters(Ghysels, Santa-Clara, and Valkanov suggested  $\boldsymbol{\theta} = (\theta_1, \theta_2)$  for all j), while in equation (2),  $\{\gamma_j\}$  are different parameters. In equation (4), we need limited number of parameters to estimate regardless of the number of lags of  $x_t$  or the value of  $(m)$ , so the parsimony principle is satisfied here[8,18].

Now, for more explanations, it is useful to introduce a simple R-MIDAS regression, and some related notations.

### **2.4 Simple R-MIDAS Regression**

Suppose that the variable  $y_t$  is available once between  $t-1$  and  $t$  (say, monthly), another variable  $x_t^{(m)}$  is observed  $(m)$  times in the same period (say, daily or  $m=22$ ), and that we are interested in the dynamic relation between  $y_t$  and  $x_t^{(m)}$ . In other words, we want to project the left-hand side variable  $y_t$  of the model equation on the history of lagged observations of  $x_{t-j/m}^{(m)}$  [2,3,4]. The latter term denotes the higher-frequency variable and its timing lags are expressed as fractions of the unit interval between  $(t-1)$  and  $(t)$ . A simple R-MIDAS model is :

$$y_t = \beta_0 + \beta_1 B\left(L^{\frac{1}{m}}; \boldsymbol{\theta}\right) x_t^{(m)} + u_t \quad (5)$$

Where :

$B\left(L^{\frac{1}{m}}; \boldsymbol{\theta}\right) = \sum_{k=0}^K B(k; \boldsymbol{\theta}) L^{\frac{k}{m}}$ ; K should be specified properly(see section 2.6 for lag specification criteria) and  $L^{\frac{k}{m}}$  is the lag operator such that :  $L^{\frac{k}{m}} x_t^{(m)} = x_{t-\left(\frac{k}{m}\right)}^{(m)}$ .

Note that the summation started from k=0 which guarantees  $x_t^{(m)}$  to be in the model, also  $x_t^{(m)}$  values have corresponding  $y_t$  values, but  $x_{t-\left(\frac{k}{m}\right)}^{(m)}$  values have corresponding missing  $y_t$  values for  $k>0$ ,  $u_t$  is white noise process[2,3,4].

the lag coefficients  $B(k; \boldsymbol{\theta})$  of the corresponding lag operator  $L^{\frac{k}{m}}$  are parameterized as function of a small-dimensional vector of parameters  $\boldsymbol{\theta}$ , usually the vector  $\boldsymbol{\theta} = (\theta_1, \theta_2)$ .

To explain simple R-MIDAS regression model (5), let us return to the quarterly-monthly variables. We have  $m=3$  and we assume that  $K=m-1=2^*$ . By equation (5), our simple R-MIDAS regression will be:

$$y_t = \beta_0 + \beta_1 \left\{ \sum_{k=0}^{K=2} B(k; \boldsymbol{\theta}) L^{\frac{k}{m}} x_t^{(m)} \right\} + u_t = \beta_0 + \beta_1 \left\{ \sum_{k=0}^{K=2} B(k; \boldsymbol{\theta}) x_{t-\left(\frac{k}{m}\right)}^{(m)} \right\} + u_t$$

Then, it can be expanded as the following:

$$y_t = \beta_0 + \beta_1 \left\{ B(0; \boldsymbol{\theta}) x_t^{(3)} + B(1; \boldsymbol{\theta}) x_{t-\left(\frac{1}{3}\right)}^{(3)} + B(2; \boldsymbol{\theta}) x_{t-\left(\frac{2}{3}\right)}^{(3)} \right\} + u_t \quad (6)$$

From equation (6), it is clear that the lower-frequency variable  $y_t$  is regressed on each of the following: 1) The higher-frequency variable  $x_t(x_t^{(3)})$  whose values represent the third month in each quarter; 2) The first lag of the higher-frequency variable  $x_t(x_{t-\left(\frac{1}{3}\right)}^{(3)})$  whose values represent the second month in each quarter, and 3)

The second lag of the higher-frequency variable  $x_t(x_{t-\left(\frac{2}{3}\right)}^{(3)})$  whose values represent the first month in each quarter. The term  $y_t$  and its lags can be added to the right hand side of this equation in order to forecast the future values( $y_{t+1}$ ) in the left hand side[2,8].

---

\* usually  $K = m-1$  such in quarterly-monthly data, in such data we have  $m=3$  so K is taken equal 2 since the summation starting from zero .

Again as we noted from equation(3), equation(6) can represent forecasting equation to the next period of  $t(t+1)$  and it does not contain leads. The lead terms can be added are:  $B^*(0; \theta)x_{t+1}^{(3)} + B^*(1; \theta)x_{t+1-\frac{1}{3}}^{(3)} + B^*(2; \theta)x_{t+1-\frac{2}{3}}^{(3)}$ . These terms will be used when monthly observations of  $(x_t)$  in the period  $(t+1)$  are available[8].

Returning to our explanation example (that was shown in table (1) before), the following table(table (2)) shows the quarterly-monthly data structure of the simple MIDAS regression model (6).

**Table (2): Two Years Quarterly-Monthly Data Structure for Eq(6)**

<b>t</b>	Quarterly-low-frequency variable $y_t$	Third-Months variable $x_t^{(3)} = x_t$	Second-Months variable $x_{t-\frac{1}{3}}^{(3)}$	First-Months variable $x_{t-\frac{2}{3}}^{(3)}$
1	$y_1$	$x_1^{(3)} = x_1$	$x_{1-\frac{1}{3}}^{(3)}$	$x_{1-\frac{2}{3}}^{(3)}$
2	$y_2$	$x_2^{(3)} = x_2$	$x_{2-\frac{1}{3}}^{(3)}$	$x_{2-\frac{2}{3}}^{(3)}$
3	$y_3$	$x_3^{(3)} = x_3$	$x_{3-\frac{1}{3}}^{(3)}$	$x_{3-\frac{2}{3}}^{(3)}$
4	$y_4$	$x_4^{(3)} = x_4$	$x_{4-\frac{1}{3}}^{(3)}$	$x_{4-\frac{2}{3}}^{(3)}$
5	$y_5$	$x_5^{(3)} = x_5$	$x_{5-\frac{1}{3}}^{(3)}$	$x_{5-\frac{2}{3}}^{(3)}$
6	$y_6$	$x_6^{(3)} = x_6$	$x_{6-\frac{1}{3}}^{(3)}$	$x_{6-\frac{2}{3}}^{(3)}$
7	$y_7$	$x_7^{(3)} = x_7$	$x_{7-\frac{1}{3}}^{(3)}$	$x_{7-\frac{2}{3}}^{(3)}$
8	$y_8$	$x_8^{(3)} = x_8$	$x_{8-\frac{1}{3}}^{(3)}$	$x_{8-\frac{2}{3}}^{(3)}$

## **2.5 Polynomial Specifications**

Parameterization of coefficients  $B(k; \theta)$  in a parsimonious fashion is one of the features of MIDAS approach[2]. We will discuss various methods of specifications of lagged coefficients  $B(k; \theta)$  of MIDAS regression models.

### **2.5.1 Exponential Almon Lag Polynomial**

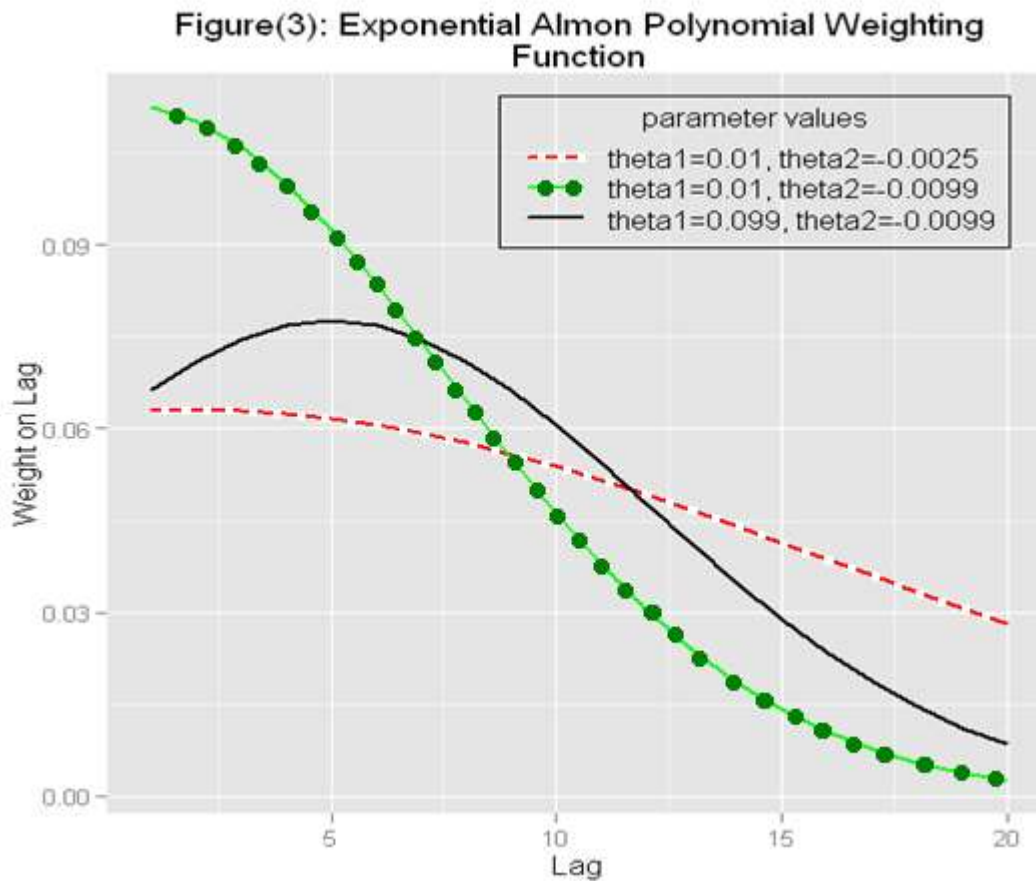
The first method for polynomial specification is to use the Exponential Almon Lag method to parameterize  $B(k; \theta)$ [2]. General Exponential Almon Lag function is defined as:

$$B(k; \boldsymbol{\theta}) = \frac{e^{\theta_1 k + \theta_2 k^2 + \dots + \theta_Q k^Q}}{\sum_{k=1}^K e^{\theta_1 k + \theta_2 k^2 + \dots + \theta_Q k^Q}}$$

Ghysels, Santa-Clara and Valkanov[1] used this functional form with two parameters, i.e.  $\boldsymbol{\theta} = (\theta_1, \theta_2)$ , because this function is known to be quite flexible, and can take various shapes with only a few parameters, so it will be in this form:

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

Figure (3) shows some shapes of the Exponential Almon Lag weighting function with different values of two parameters  $\theta_1$  and  $\theta_2$ . We note that different values of parameters obtain decreasing or hump-shaped weighting functions. It is clear from the figure(3) that later lags of the higher-frequency variable  $x_t$  are assigned to a lower weights[14].





We have some notes and properties about this function:

1. When  $\theta_1 = \theta_2 = 0$ , then  $B(k; \boldsymbol{\theta})$  represents the same weights for all  $k$ . In the case of quarterly-monthly data,  $K=m=3$ , and  $B(k; \boldsymbol{\theta}) = \frac{1}{3}$  for all  $k$ , which is the same as the Time-Averaging Aggregation.
2. The sum of  $B(k; \boldsymbol{\theta})$  is up to one, i.e.  $\sum_{k=1}^K B(k; \boldsymbol{\theta}) = 1$ .
3. Since  $B(k; \boldsymbol{\theta})$  is a nonlinear function of parameters  $\theta_1$  and  $\theta_2$ , these parameters can be estimated by the nonlinear least squares method (NLS)[2,4,8,32].

### **2.5.2 Beta Lag function**

The second polynomial specification is to use the Beta function, so-called as Beta Lag Function which has only two parameters such that:

$$B(k; \theta_1, \theta_2) = \frac{f\left(\frac{k}{K}; \theta_1, \theta_2\right)}{\sum_{k=0}^K f\left(\frac{k}{K}; \theta_1, \theta_2\right)}$$

Where:

$$f(x; \theta_1, \theta_2) = \frac{x^{\theta_1-1}(1-x)^{\theta_2-1}\Gamma(\theta_1+\theta_2)}{\Gamma(\theta_1)\Gamma(\theta_2)}, \quad \Gamma(\theta) = \int_0^\infty e^{-x} x^{\theta-1} dx.$$

The Beta lag Function has the same properties stated for the Exponential Almon Lag function except that the Time-Averaging Aggregation is obtained by the Beta lag Function when  $\theta_1 = \theta_2 = 1$ [2,4,8].

### **2.5.3 Almon Lag Polynomial of order (P)**

The almon lag polynomial is used to specify  $\beta_1 B(k; \boldsymbol{\theta})$  **together or jointly** in MIDAS regression model (6) which is defined as:

$$\beta_1 B(k; \theta_0, \theta_1, \dots, \theta_p) = \theta_0 + \sum_{p=0}^P \theta_p k^p$$

This polynomial can be written in matrix form as:

$$\beta_1 \begin{bmatrix} B(0; \boldsymbol{\theta}) \\ B(1; \boldsymbol{\theta}) \\ B(2; \boldsymbol{\theta}) \\ \vdots \\ B(k; \boldsymbol{\theta}) \\ \vdots \\ B(K; \boldsymbol{\theta}) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0^2 & \dots & 0^P \\ 1 & 1 & 1^2 & \dots & 1^P \\ 1 & 2 & 2^2 & \dots & 2^P \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & k & k^2 & \dots & k^P \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & K & K^2 & \dots & K^P \end{bmatrix} \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_P \end{bmatrix}$$

For example, by this polynomial, we have the following:

$$\begin{aligned}\beta_1 B(0; \boldsymbol{\theta}) &= \theta_0 \\ \beta_1 B(1; \boldsymbol{\theta}) &= \theta_0 + \theta_1 + \theta_2 + \dots + \theta_p \\ \beta_1 B(2; \boldsymbol{\theta}) &= \theta_0 + 2\theta_1 + 4\theta_2 + 8\theta_3 + \dots + 2^p \theta_p \\ \beta_1 B(3; \boldsymbol{\theta}) &= \theta_0 + 3\theta_1 + 9\theta_2 + 27\theta_3 + \dots + 3^p \theta_p \\ &\vdots \\ &\vdots \\ &\vdots \\ \beta_1 B(K; \boldsymbol{\theta}) &= \theta_0 + K\theta_1 + K^2\theta_2 + K^3\theta_3 + \dots + K^p \theta_p\end{aligned}$$

For more clarification, let us return to MIDAS model (6) with vector  $\boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2)$ . Using weights of Almon-Lag-Polynomial of order (2), assuming  $K=m-1=2$ , equation (6) becomes:

$$y_t = \beta_0 + \theta_0 x_t^{(3)} + (\theta_0 + \theta_1 + \theta_2) x_{t-\frac{1}{3}}^{(3)} + (\theta_0 + 2\theta_1 + 4\theta_2) x_{t-\frac{2}{3}}^{(3)} + u_t \quad (7)$$

Where  $\beta_1 B(0; \boldsymbol{\theta}) = \theta_0$ ,  $\beta_1 B(1; \boldsymbol{\theta}) = \theta_0 + \theta_1 + \theta_2$ ,  $\beta_1 B(2; \boldsymbol{\theta}) = \theta_0 + 2\theta_1 + 4\theta_2$ . Arranging equation(7) will give the following equation:

$$y_t = \beta_0 + \theta_0 x'_{t1} + \theta_1 x'_{t2} + \theta_2 x'_{t3} + u_t^{(3)} \quad (8)$$

Where :  $x'_{t1} = x_t^{(3)} + x_{t-\frac{1}{3}}^{(3)} + x_{t-\frac{2}{3}}^{(3)}$ ,  $x'_{t2} = x_{t-\frac{1}{3}}^{(3)} + 2x_{t-\frac{2}{3}}^{(3)}$  and

$$x'_{t3} = x_{t-\frac{1}{3}}^{(3)} + 4x_{t-\frac{2}{3}}^{(3)}$$

Now, the parameters  $\theta_0, \theta_1$  and  $\theta_2$  could be estimated by ordinary least squares method(OLS) such that :

$$\boldsymbol{\theta} = (\mathbf{X}^T \mathbf{X})^{-1} (\mathbf{X}^T \mathbf{Y})$$

Where The matrix  $\mathbf{X} = \{1, x'_{t1}, x'_{t2}, x'_{t3}\}$ ,  $\boldsymbol{\theta} = (\beta_0, \theta_0, \theta_1, \theta_2)$ ,  $\mathbf{Y} = \{y_t\}$ , the symbol ( $^T$ ) here means matrix transposition[2,15,16].

Ghysels, Santa-Clara, and Valkanov (2004) suggest that MIDAS models can be estimated under general conditions via non-linear least squares (NLS), quasi-maximum-likelihood (MLE) or general method of moments (GMM)(see section 2.10)[4]. The other specification is to define MIDAS regression and estimating parameters with step-function method[2,15,16].

## **2.6 Lag Length Selection Criteria**

As we saw in previous sections, there are number of lags selected to include in MIDAS model, i.e. the MIDAS regression model depends on  $K$  which is the maximum effective lag of the higher-frequency variable  $x$  should be taken. For simplicity, we have selected  $K=m$  in our explanation examples, but the determination of lag length for any time series is very important issue in all studies. In the literature of linear regression, there are various selection criteria used to determine the best subset of independent variables should be included in the model. Of course we always interested to get the model with minimum mean square error(MSE), but we can make the MSE smaller by adding another MA or AR terms in time series model(the same thing in linear regression). But this make a problem of getting complicated models with too many parameters(violation of parsimony principle), so the use of information criterion techniques would help us to construct models that fit the data well without having too many parameters[25].

The same idea is applied in time series models. We have various lag length selection criteria, such as: the Aikake's information criterion(AIC), Schwarz information criterion(SIC), Hannan-Quinn criterion(HQC), Final Prediction Error(FPE) and the Bayesian information criterion(BIC). Both AIC and BIC are the most popular and they have been used widely by researchers[23,24,25].

Ghysels, Santa-Clara, and Valkanov(2004) stated that standard selection procedures such as the Akaike or Schwarz criterion can be applied to select the optimal number of lags that should be included in MIDAS regression. Ghysels, Claudia, Klaus, Andreou and Kourtellos were concentrated on AIC and BIC. They also compared between the two criteria to get the optimal lag length and to improve results of MIDAS models[3,4]. In BIC which used much more, we choose the model that has the minimum value of:

$$BIC_M = \log(\hat{\sigma}^2) + \frac{M}{T} \log(T) \quad (9)$$

Where  $\hat{\sigma}^2$  is the estimated error-variance after fitting the model(usually the sum of squared residuals divided by the number of observations( $T$ )),  $M$  is the number of estimated parameters and  $T$  is the length of the time series(low-frequency variable length) [4,20].

## **2.7 Testing Unit Root and Cointegration**

### **2.7.1 Unit Root**

Time series data consist of observations, which are considered as a realization of random variables that can be described by some stochastic processes. The concept of stationarity is related to the properties of these stochastic processes. The concept of weak stationarity means that the data are assumed to be stationary if the means, variances and covariances of the series are independent of time, rather than the entire distribution. Nonstationarity in a time series occurs when there is no constant mean, no constant variance or both of these properties. It can originate from various sources but the most important one is the unit root [54,42,43].

Any sequence that contains one or more characteristic roots that are equal to one is called a unit root process. The simplest model that may contain a unit root is the AR(1) model. Consider the autoregressive process of order one, AR(1), below:

$$y_t = \phi y_{t-1} + u_t$$

Where  $u_t$  denotes a serially uncorrected white noise error term with a mean of zero and a constant variance. If  $\phi = 1$ , the above equation becomes a random walk without drift model, that is, a nonstationary process. When this happens, we face what is known as the unit root problem. This means that we are faced with a situation of nonstationarity in the series. If, however,  $\phi < 1$ , then the series  $y_t$  is stationary. The stationarity of the series is important because correlation could persist in nonstationary time series even if the sample is very large and may result in what is called spurious (or nonsense) regression. The unit root problem can be solved, or stationarity can be achieved, by differencing the data set [54,42,43].

We have many methods to stationarity, but the most famous one is The augmented Dickey-Fuller (ADF) test, The basic idea behind the ADF unit root test for nonstationarity is to simply regress  $y_t$  on its (one period) lagged value  $y_{t-1}$  and find out if the estimated  $\phi$  is statistically equal to 1 or not. The above AR(1) equation can be manipulated by subtracting  $y_{t-1}$  from both sides to obtain

$$y_t - y_{t-1} = (\phi - 1)y_{t-1} + u_t$$

Which can be written as

$$\Delta y_t = \delta y_{t-1} + u_t$$

Where  $\delta = \phi - 1$ , and  $\Delta$  is the first difference operator.

In practice, we shall estimate the last equation and test for the null hypothesis of  $\delta = 0$  against the alternative of  $\delta \neq 0$ . If  $\delta = 0$ , then  $\phi = 1$ , meaning that we have a unit root problem and the series under consideration is nonstationary. It should be noted that under the null hypothesis  $\delta = 0$ , the t-value of the estimated coefficient of  $y_{t-1}$  does not follow the t-distribution even in large samples. This means that the t-value does not have an asymptotic normal distribution. The decision to reject or not to reject the null hypothesis of  $\delta = 0$  is based on the Dickey-Fuller (DF) critical values of the  $\tau$ (tau) statistic(Dickey-Fuller t-distribution). The DF test is based on an assumption that the error terms  $u_t$  are uncorrelated[54,42,43].

However, in practice, the error term in the DF test usually show evidence of serial correlation. To solve this problem, Dickey and Fuller have developed a test know as the Augmented Dickey-Fuller (ADF) test. In the ADF test, the lags of the first difference are included in the regression equation in order to make the error term  $u_t$  white noise and, therefore, the regression equation is presented in the following form:

$$\Delta y_t = \delta y_{t-1} + \sum_{i=1}^m \alpha_i \Delta y_{t-i} + u_t$$

To be more specific, the intercept may be included, as well as a time trend  $t$ , after which the model becomes

$$\Delta y_t = \beta_0 + \beta_1 t + \delta y_{t-1} + \sum_{i=1}^m \alpha_i \Delta y_{t-i} + u_t$$

where  $\beta_0$  is a constant,  $\beta_1$  the coefficient on a time trend series,  $\delta$  the coefficient of  $y_{t-1}$ ,  $m$  is the lag order of the autoregressive process,  $\Delta y_t = y_t - y_{t-1}$  are first difference of  $y_t$ ,  $y_{t-1}$  are lagged values of order one of  $y_t$ ,  $\Delta y_{t-i}$  are changes in lagged values, and  $u_t$  is the white noise. The test procedure for unit roots is similar to statistical tests for hypothesis, that is: (i) Set the null and alternative hypothesis as  $H_0: \delta = 0$  against  $H_1: \delta < 0$ . (ii) Determine the test statistic using  $F_\tau = \frac{\hat{\delta}}{SE(\hat{\delta})}$  where  $SE(\hat{\delta})$  is the standard error of  $\hat{\delta}$ . (iii) Compare the calculated test statistic  $F_\tau$  with the critical value from Dickey-Fuller table to reject or not to reject the null hypothesis. (iv) The ADF test is a lower-tailed test, so if  $F_\tau$  is less than the critical value, then the null hypothesis of unit root is rejected and the conclusion is that the variable of the series does not contain a unit root and is stationary[54,42,43].

The DF and ADF tests are similar since they have the same asymptotic distribution. Finally, there are numerous unit root tests, such as the Phillips-Perron test and the Schmidt-Phillips test, but the most notable and commonly used is the ADF test, which will be used in our study[54,42,43].

### **2.7.2 Cointegration**

It is possible for two integrated series to move together in a nonstationary way, so that their difference or any other linear combination is stationary. These series are said to be cointegrated. Stationarity is like a rubber band pulling a series back to the fixed mean[51,52,54]. Cointegration is like a rubber band pulling the two series back to a fixed relationship with each other, even though both series are not pulled back to a fixed mean. If  $y$  and  $x$  are both integrated, we cannot rely on OLS standard errors or  $t$  statistics. By differencing, we can avoid spurious regressions:

If  $y_t = \beta_0 + \beta_1 x_t + u_t$  then  $\Delta y_t = \Delta x_t + \Delta u_t$ .  $\Delta u_t$  is stationary as long as  $u_t$  is  $I(0)$  or  $I(1)$ . Note that the differenced equation has no history. Now let us Suppose that  $u_t$  is  $I(1)$ , This means that the difference  $u_t = y_t - \beta_0 - \beta_1 x_t$  is not mean-reverting and there is no long-run tendency for  $y$  to stay in the fixed relationship with  $x$ , so there is no cointegration between  $y$  and  $x$ , in this case  $\Delta u_t$  is  $I(0)$ , if  $y_t$  is high relative to  $x_t$  due to a large positive  $u_t$ , then there is no tendency for  $y$  to come back to  $x$  after  $t$ , thus estimation of differenced equation is appropriate[51,52,54].

Now Suppose that  $u_t$  is  $I(0)$ , that means that the levels of  $y$  and  $x$  tend to stay close to the relationship given by the equation. Suppose that there is a large positive  $u_t$  that puts  $y_t$  about its long-run equilibrium level in relation to  $x_t$ . With stationary  $u_t$ , we expect the level of  $y$  to return to the long-run relationship with  $x$  over time, and stationarity of  $u_t$  implies that  $\text{corr}(u_t, u_{t+s}) \rightarrow 0$  as  $s \rightarrow \infty$ . Thus, future values of  $\Delta y$  should tend to be smaller than those predicted by  $\Delta x$  in order to close the gap. In terms of the error terms, a large positive  $u_t$  should be followed by negative  $\Delta u_t$  values to return  $u_t$  to zero if  $u_t$  is stationary, and this is the situation where  $y$  and  $x$  are cointegrated. This is not reflected in the differenced equation, which says that future values of  $\Delta y$  are only related to the future  $\Delta x$  values, which means there is no tendency to eliminate the gap that opened up at  $t$  [51,52,54].

In the cointegrated case, If we estimate in differences, we are missing the history of knowing how  $y$  will be pulled back into its long-run relationship with  $x$ . If we estimate in the original levels, we cannot rely on our test statistics because the variables (though not the error term) are nonstationary[51,52,54].

The appropriate model for the cointegrated case is the error-correction model of Hendry and Sargan, ECM consists of two equations:

The long-run cointegrating equation:  $y_t = \theta_0 + \theta_1 x_t + u_t$ , where  $u_t$  is  $I(0)$ .

The short-run (ECM) adjustment equation:

$$\Delta y_t = \beta_0 + \sum_{i=1}^p \beta_i L^i \Delta y_t + \sum_{i=1}^p \gamma_i L^i \Delta x_t + \alpha(y_{t-1} - \theta_0 - \theta_1 x_{t-1}) + v_t$$

Note the presence of the error-correction term with coefficient  $\alpha$  in the ECM equation, this term reflects the distance that  $y_{t-1}$  is from its long-run relationship with  $x_{t-1}$ . If  $\alpha < 0$ , then  $y_{t-1}$  above its long-run level will cause  $\Delta y_t$  to be negative when other factors are held constant, then it pulling  $y$  back toward its long-run relationship with  $x$ . Because both  $y$  and  $x$  are  $I(1)$ , their differences are  $I(0)$ . Because they are cointegrated with cointegrating vector  $\theta_0, \theta_1$ , the difference in the error correction term is also  $I(0)$ . It would not be if they weren't cointegrated and the ECM regression would be invalid. The ECM equation can be estimated by OLS without undue difficulty because all the variables are stationary[51,52,54].

In multivariate cointegration, the concept of cointegration extends to multiple variables. With more than two variables, there can be more than one cointegrating relationship (vector). Vector error-correction models (VECM) allow for the estimation of error correction regressions with multiple cointegrating vectors. In order to test for cointegration, the earliest test is Engle and Granger's extension of the ADF(Augmented Dickey-Fuller) test has two steps: 1) Regressing the cointegrating regression by OLS, 2) Testing the residuals with an ADF test. Other, more popular tests include the Johansen-Juselius test, which generalizes easily to multiple variables and multiple cointegrating relationships[51,52,53,54].

### **2.7.3 Johansen's procedure**

Johansen's procedure builds cointegrated variables directly on maximum likelihood estimation instead of relying on OLS estimation. This procedure relies heavily on the relationship between the rank of a matrix and its characteristic roots. Johansen derived the maximum likelihood estimation using sequential tests for determining the number of cointegrating vectors[54]. His method can be seen as a secondary generation approach in the sense that it builds directly on maximum likelihood instead of partly relying on least squares. In fact, Johansen's procedure is nothing more than a multivariate generalization of the Dickey-Fuller test. Consequently, he proposes two different likelihood ratio tests namely: the trace test and the maximum eigenvalue test. This procedure is a vector cointegration test

method. It has the advantage over the Engle-Granger method in that it can estimate more than one cointegration relationship, if the data set contains two or more time series[52,53,54].

Johansen's method takes as a starting point the vector autoregression (VAR) of order  $p$  given by

$$\mathbf{X}_t = \boldsymbol{\pi}_1 \mathbf{X}_{t-1} + \boldsymbol{\pi}_2 \mathbf{X}_{t-2} + \cdots + \boldsymbol{\pi}_p \mathbf{X}_{t-p} + \mathbf{U}_t$$

where  $\mathbf{X}_t$  is an  $n \times 1$  vector of variables that are integrated of order one, that is,  $I(1)$ .  $\mathbf{U}_t$  is an  $n \times 1$  vector of innovations while  $\boldsymbol{\pi}_1$  through  $\boldsymbol{\pi}_p$  are  $m \times m$  coefficient matrices. Reparameterizing the equation above, that is, subtracting  $\mathbf{X}_{t-1}$  on both sides, leads to

$$\Delta \mathbf{X}_t = \boldsymbol{\Gamma}_1 \Delta \mathbf{X}_{t-1} + \boldsymbol{\Gamma}_2 \Delta \mathbf{X}_{t-2} + \cdots + \boldsymbol{\Gamma}_{p-1} \Delta \mathbf{X}_{t-p+1} - \boldsymbol{\pi} \mathbf{X}_{t-p} + \mathbf{U}_t$$

Where  $\boldsymbol{\Gamma}_1 = \boldsymbol{\pi}_1 - \mathbf{I}$ ,  $\boldsymbol{\Gamma}_2 = \boldsymbol{\pi}_2 - \boldsymbol{\Gamma}_1$ ,  $\boldsymbol{\Gamma}_3 = \boldsymbol{\pi}_3 - \boldsymbol{\Gamma}_2$ , ... ,  $\boldsymbol{\Gamma}_{p-1} = \boldsymbol{\pi}_{p-1} - \boldsymbol{\Gamma}_{p-2}$ , and  $\boldsymbol{\pi} = \mathbf{I} - \boldsymbol{\pi}_1 - \boldsymbol{\pi}_2 - \cdots - \boldsymbol{\pi}_p$ . The matrix  $\boldsymbol{\pi}$  determines the extent to which the system is cointegrated and is called the impact matrix[54].

Returning to the last reparameterised equation, if we consider the first equation of the system as:

$$\Delta X_{1,t} = \gamma'_{11} \Delta X_{1,t-1} + \gamma'_{12} \Delta X_{1,t-2} + \cdots + \gamma'_{1p-1} \Delta X_{1,t-p+1} - \pi'_1 X_{1,t-p} + u_{1t}$$

where  $\gamma'_{ij}$  is the first row of  $\boldsymbol{\Gamma}_j$ ,  $j = 1, 2, \dots, p-1$ , and  $\pi'_1$  is the first row  $\boldsymbol{\pi}$ .

Here  $\Delta X_{1,t}$  is stationary, that is,  $I(0)$ ,  $j = 1, 2, \dots, p-1$  are all  $I(0)$ ,  $u_t$  is assumed to be  $I(0)$  and so for a meaningful equation,  $\pi'_1 X_{1,t-p}$  must be stationary,  $I(0)$ .

If none of the components of  $\mathbf{X}_t$  are cointegrated, they must be zero. On the other hand, if they are cointegrated, all the rows of  $\boldsymbol{\pi}$  must be cointegrated but not necessarily distinct. This is because the number of distinct cointegrating vectors depends on the row rank of  $\boldsymbol{\pi}$ [52].

The matrix  $\boldsymbol{\pi}$  is of order  $m \times m$ . If it has rank  $m$ , that is,  $m$  number of linearly independent rows or columns, then it forms a basis for  $m$ -dimensional vector space. This implies that all  $m \times 1$  vectors can be generated as linear combinations of its row. Any of these linear combination of the rows would lead to stationarity, meaning that  $\mathbf{X}_{t-p}$  has stationary components if the rank of  $\boldsymbol{\pi}$  is  $r < m$ [52].

We may write  $\boldsymbol{\pi} = \underline{\boldsymbol{\beta}} \underline{\boldsymbol{\alpha}}'$  for suitable  $m \times r$  matrices,  $\underline{\boldsymbol{\beta}}$  and  $\underline{\boldsymbol{\alpha}}$  Here



$$\underline{\alpha}' = \begin{bmatrix} \alpha'_1 \\ \alpha'_2 \\ \vdots \\ \alpha'_r \end{bmatrix}, \quad \underline{\beta} = [\beta_1, \beta_2, \dots, \beta_r], \quad \text{then } \underline{\pi} \mathbf{X}_{t-p} = \underline{\beta} \underline{\alpha}' \mathbf{X}_{t-p}, \quad \text{and all linear}$$

combinations of  $\underline{\alpha}' \mathbf{X}_{t-p}$  are stationary. It should be noted that we have to perform the ADF test to assess the order of integration of each variable before applying Johansen's procedure. Johansen's procedure estimates the VAR subject to  $\underline{\pi} = \underline{\beta} \underline{\alpha}'$  for various values of  $r$  number of cointegrating vectors, using the maximum likelihood estimator assuming  $u_t \sim iidN(0, \Sigma)$ , His estimate can thus be rewritten as

$$\Delta \mathbf{X}_t = \Gamma_1 \Delta \mathbf{X}_{t-1} + \Gamma_2 \Delta \mathbf{X}_{t-2} + \dots + \Gamma_{p-1} \Delta \mathbf{X}_{t-p+1} - \underline{\beta} \underline{\alpha}' \mathbf{X}_{t-p} + \mathbf{U}_t$$

To detect the number of cointegrating vectors, Johansen proposed two likelihood ratio tests namely:

- The trace test
- The maximum eigenvalue.

The trace test tests the null hypothesis of  $r$  cointegrating vectors against the alternative hypothesis of  $n$  cointegrating vectors, If  $r = 0$ , it means that there is no relationship among the variables that is stationary. The test statistic is given by

$$J_{trace} = -T \sum_{i=r+1}^n \ln(1 - \hat{\lambda}_i)$$

The maximum eigenvalue test, on the other hand, tests the null hypothesis of  $r$  cointegrating vectors against the alternative hypothesis of  $(r+1)$  cointegrating vectors. Its test statistic is given by

$$J_{max} = -T(1 - \hat{\lambda}_{r+1})$$

Where  $T$  is the sample size, and  $\hat{\lambda}_i$  is the  $i$ 'th largest canonical correlation. For example, if the null hypothesis of no cointegration ( $r = 0$ ) against the alternative of presence of one or more cointegrating vector is rejected at specific level of significance say (10%) of significance in both techniques (trace test and maximum eigenvalue); this implies that cointegration exists between series data set. If the null hypothesis ( $r \leq 1$ ) and ( $r \leq 2$ ) against the alternative of the existence of two or three cointegrating vectors is not rejected by both tests; This means that there is no more than one cointegration relationship between series data set(see for example Table A2 in Appendix A)[52,53,54].

The method of Johansen assumes that the cointegrating vector remains constant during the period of study. In reality, it is possible that the long-run relationships between the underlying variables change. The reason for this might be technological progress, economic crisis, changes in people's preferences and behavior accordingly, policy or regime alteration and institutional development. This is especially the case if the sample period is long[52,53,54].

## **2.8 Cointegration-MIDAS and ECM-MIDAS models**

In almost all the literature, MIDAS regressions are applied to stationary time series or to transformed non-stationary variables by the first differences. Goetz, Hecq and Urbain [21] worked with mixed frequency time series that are integrated of order one  $I(1)$  and possibly cointegrated. They introduced a mixed-frequency autoregressive distributed lag ADL(see section 2.9 about ADL) model and they derived the cointegrating long-run relationships, then they plugged it in the error correction model ECM to derive the short-run relationships. Finally, they compared between number of approaches by simulation methods in order to know which one has the best forecasting performance. We interested on the following approaches that they studied: unrestricted MIDAS regressions(U-MIDAS), restricted MIDAS regressions(R-MIDAS) and Time-Averaging regressions. The estimated forecasting ECM models that they used relative to each one of these approaches can be written as the following:

$$\widehat{\Delta y}_t = \widehat{c} + \widehat{\gamma} \widehat{u}_{t-1} + \widehat{\alpha} y_{t-1} + \sum_{k=0}^{m-1} \widehat{B}_k \Delta^{(m)} x_{t-\left(\frac{k}{m}\right)}^{(m)} \quad \text{(U- MIDAS)} \quad (10)$$

$$\widehat{\Delta y}_t = \widehat{c} + \widehat{\gamma} \widehat{u}_{t-1} + \widehat{\alpha} y_{t-1} + \widehat{\beta} \sum_{k=0}^{m-1} B_k(\widehat{\theta}_1, \widehat{\theta}_2) \Delta^{(m)} x_{t-\left(\frac{k}{m}\right)}^{(m)} \quad \text{(R- MIDAS)} \quad (11)$$

$$\widehat{\Delta y}_t = \widehat{c} + \widehat{\gamma} \widehat{u}_{t-1} + \widehat{\alpha} y_{t-1} + \widehat{\beta} \Delta \left\{ \frac{1}{m} \sum_{k=0}^{m-1} x_{t-\left(\frac{k}{m}\right)}^{(m)} \right\} \quad \text{(Time-Averaging)} \quad (12)$$

Where :  $c$  is the constant term, the symbol  $\Delta$  means the first difference for the low-frequency variable, the symbol  $\Delta^{(m)}$  means the first difference for the high-frequency variable,  $u_{t-1}$  is the first lag of the residuals resulting from the long-run cointegrating relationship and  $\widehat{u}_{t-1}$  is its estimator(under assumption that  $u_t$  is white noise process).

Goetz, Hecq and Urbain[21] showed that the ECM built by MIDAS regressions has the best forecasting performance. They employed Exponential Almon-Lag polynomial to estimate the weights  $B_i(\widehat{\theta}_1, \widehat{\theta}_2)$  with initial value of  $\theta_2 \leq 0$ , because it is extremely flexible and it has various possible shapes. The Exponential Almon-Lag

weights make estimation to be completely data driven, and it allows to include variables in both short-run and long-run models more than what theoretically necessary, because the redundant variables will be assigned to zero weight(since the sum of all Exponential Almon Lag weights is 1).

Also, they concluded that when building the short-run ECM, it is better to use  $(m)$  terms than more or less than  $(m)$  terms; where  $m$  is the number of the high-frequency observations between the period  $(t)$  and  $(t-1)$  of the low-frequency variable.

We chose to use MIDAS regression **(10)** or **(11)** to construct the short-run ECM with Exponential Almon Lag weights assigned to  $(m)$  terms(or lags) of the high-frequency variables[21]. Also it is suitable to use MIDAS regression to construct the long-run relationship which will produce the imposed  $u_{t-1}$  terms by the following ADL-RMIDAS equation:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 \sum_{k=0}^{m-1} B_k(\theta_1, \theta_2) x_{t-\left(\frac{k}{m}\right)}^{(m)} + u_t \quad (13)$$

equation **(5)** with  $K=m-1$ \*, and the autoregressive terms  $(y_{t-1})$  can be stay or dropped if or if not necessary from the above equations **(10\_13)**. So the equilibrium terms  $\hat{u}_{t-1}$  can be obtained by the equation :

$$\hat{u}_{t-1} = y_{t-1} - \hat{\beta}_0 - \hat{\beta}_1 y_{t-2} - \hat{\beta}_2 \sum_{k=0}^{m-1} B_k(\hat{\theta}_1, \hat{\theta}_2) x_{t-1-\left(\frac{k}{m}\right)}^{(m)} \quad (14)$$

because from equation**(13)**, we have

$u_t = y_t - \beta_0 - \beta_1 y_{t-1} - \beta_2 \sum_{k=0}^{m-1} B_k(\theta_1, \theta_2) x_{t-\left(\frac{k}{m}\right)}^{(m)}$ . Finally, the ADL-UMIDAS regression will stay in mind to use if necessary[21,28].

## **2.9 Distributed-Lag Models**

Univariate time-series models are an interesting and useful building block, but we are almost always interested not just in  $y$ 's behavior by itself but also in how  $x$  affects  $y$ . In time-series models, this effect is often dynamic(spread out over time). Thus we can define a ***distributed-lag model*** as a dynamic model in which the effect of a regressor  $x$  on  $y$  occurs over time rather than all at once[51].

---

\* The R-package program splits the high-frequency variable and name each split as X0/3, X1/3, X2/3 in quarterly-monthly data, for example X0/3 means the series consisted of the last month per quarter, similarly X1/3 means the series consisted of the second month per quarter, and so forth.

the dynamic effect of  $x$  on  $y$  involves not just  $\frac{\partial y}{\partial x}$ , but a more complex set of dynamic multipliers:  $\frac{\partial y_t}{\partial x_t}$ ,  $\frac{\partial y_{t+1}}{\partial x_t}$ ,  $\frac{\partial y_{t+2}}{\partial x_t}$ , etc, so we estimate the dynamic effects of  $x$  on  $y$  with distributed-lag models[51].

In general, the distributed-lag model has the form

$$y_t = \alpha + \sum_{i=0}^{\infty} \beta_i x_{t-i} + u_t$$

But of course, we cannot estimate an infinite number of lag coefficients  $\beta_i$ , so we must either truncate or find another way to approximate an infinite lag structure. We can easily have additional regressors with either the same or different lag structures. When the distributed-lag model above is truncated, then we have the Finite distributed lag model

$$y_t = \alpha + \sum_{i=0}^r \beta_i x_{t-i} + u_t$$

The finite distributed lag model can be estimated by OLS. The dynamic multipliers in this case are  $\frac{\partial y_t}{\partial x_{t-s}} = \beta_s$ ,  $s=0,1,\dots,r$  and zero other wise. The cumulative dynamic multipliers which denotes to the effect of permanent change in  $x$  are  $\sum_{i=0}^s \beta_i$  [51].

To see the interpretation of the lag weights, we consider two special cases: a temporary change in  $x$  and a permanent change in  $x$ . For the temporary change, the change in  $x$  is zero except in period  $t$ , where it is one. Here we Suppose that  $x$  changes temporarily by one unit in period  $t$ , then returns to its original level for periods  $t + 1$  and all future periods; in period  $t$ ,  $x_t$  is changed by one unit but all earlier  $x$  values are unchanged. In this case,  $y_t$  increases by  $\beta_0$  when  $x_t$  changes by one with  $x_{t-1}, x_{t-2}, \dots, x_{t-r}$  and the disturbance term  $u_t$  unchanged. We call this immediate effect =  $\beta_0$  the *impact effect of  $x$  on  $y$*  [51].

Given that  $x_{t+1}$ ,  $x_{t-1}$ , and the disturbance are unchanged, the change in  $y$  in period  $t + 1$  is the coefficient on the first lag of  $x$  ( $\beta_1$ ). This is the *dynamic marginal effect of  $x$  on  $y$  at one lag*. By similar analysis, we can see that the effect of the temporary change in  $x$  at time  $t$  on  $y_{t+2}$  is  $\beta_2$ , and so forth. We conclude that the pattern of dynamic marginal effects of a temporary change in  $x$  on  $y$  is given by the coefficients of the lag distribution  $\beta$ 's [51].

Now let us consider the case of a permanent increase in  $x$  at time  $t$ :  $x$  increases by one unit in period  $t$  and remains higher in all periods after  $t$  than it was before  $t$ . We see from our finite distributed lag model that the effect of the change in  $x_t$  on  $y_t$  is  $\beta_0$ , because  $x_{t-1}, x_{t-2}, \dots, x_{t-r}$  and  $u_t$  are assumed unchanged. Moving forward to period  $t + 1$ , the analysis of the model equation is different from the temporary change because now  $x_{t+1}$  as well as  $x_t$  is increased by one unit. The cumulative effect on  $y_{t+1}$  is  $\beta_0 + \beta_1$ . We call this the *cumulate effect of  $x$  on  $y$  at one lag*. At two lags, we advance the equation one more period forward and now the effects of the increased levels of  $x_{t+2}, x_{t+1}$ , and  $x_t$  raising  $y_{t+2}$  by  $\beta_0 + \beta_1 + \beta_2$  units[51].

Moving ( $k > r$ ) periods into the future, the effect on  $y_{t+k}$  will be the same as the effect on  $y_{t+r}$  because once again all of the  $x$  terms on the right hand side are increased by one unit and the cumulative effect staying at  $\sum_{i=0}^r \beta_i$  for all lags starting at  $k$ . The limit of the cumulative effect as the lag length goes to infinity is called the *long-run cumulative effect of  $x$  on  $y$* . It measures how much  $y$  will eventually change in response to a permanent change in  $x$  [51].

The pattern of the dynamic marginal effects and cumulative effects tells us about both the magnitude and the timing of the effect of  $x$  on  $y$ . In the above, permanent increases in  $x$  lead to permanent increases in  $y$  that get larger over the first  $r+1$  periods of the change. Temporary changes in  $x$ , by contrast, lead to temporary changes in  $y$  that die away after  $r+1$  periods[51].

Another pattern that is plausible for some economic relationships is that permanent changes in  $x$  may lead to only temporary changes in  $y$ . For example, standard macroeconomic theory tells us that changes in the rate of monetary growth have only temporary effects on real output growth. In such a situation, the positive marginal effects at short lags ( $\beta_0, \beta_1$ , and  $\beta_2$ , perhaps) would be offset by negative marginal effects at longer lags so that the long-run cumulative effect (the sum of all the  $\beta$  coefficients) is zero[51].

Is it possible to have a dynamic relationship in which temporary changes in  $x$  lead to permanent changes in  $y$ , but this creates difficulty for an equation. For a temporary change in  $x$  to have a permanent effect on  $y$ , the  $\beta$  coefficients in the infinite DL model could not approach zero as  $i$  gets large. This obviously rules out any finite lag distribution, and indeed creates trouble for any specification because the dynamic relationship between the two variables is effectively non-stationary. To cope with this situation, it is usually best to redefine either  $x$  or  $y$  to eliminate the problem[51].

In specifying dynamic econometric models, it is crucial to think very carefully about the nature of the dynamic relationship among the variables. We must decide how we would expect  $y$  to respond over time to a one-time change in  $x$ , then define the variables as levels or changes in order to represent the expected relationship with a stationary lag function[51].

The autoregressive component of the ARMA model involves using one or more lagged values of  $y$  as determinants of the current value  $y_t$ . We can apply the same method in a distributed-lag context by adding  $y_{t-1}$  and possibly additional lags to the right-hand side. The simplest model is the **Koyck lag**, which has one lag of  $y$  on the right-hand side with only the current value of  $x$  [51].

The first order autoregressive Koyck lag model with regressors is:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \delta_0 x_t + u_t$$

By this model we have  $\frac{\partial y_t}{\partial x_t} = \delta_0$ ,  $\frac{\partial y_{t+1}}{\partial x_t} = \beta_1 \delta_0$ , ...,  $\frac{\partial y_{t+s}}{\partial x_t} = \beta_1^s \delta_0$ . Thus, dynamic multipliers start at  $\delta_0$  and decay exponentially to zero over infinite time. Thus, this is effectively a distributed lag of infinite length, but with only 2 parameters (plus intercept) to estimate. Cumulative multipliers are  $\sum_{i=0}^s \frac{\partial y_{t+i}}{\partial x_t} = \delta_0 \sum_{i=0}^s \beta_1^i$ , the long-run effect of a permanent change is  $\delta_0 \sum_{i=0}^{\infty} \beta_1^i = \frac{\delta_0}{1-\beta_1}$ . Estimation of this model has the potential problem of inconsistency if  $u_t$  is serially correlated, this is a serious problem, especially as some of the test statistics for serial correlation of the error are biased when the lagged dependent variable is present. Koyck lag is parsimonious and fits lots of lagged relationships well, With multiple regressors, the Koyck lag applies the same lag structure (rate of decay) to all regressors[51].

We can generalize the Koyck lag model to longer lags:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \delta_0 x_t + u_t$$

Dynamic multipliers for this model are determined by coefficients of infinite lag polynomial  $[\sum_{i=1}^p \beta_i L^i]^{-1} = [\beta(L)]^{-1}$ , where  $L$  denotes to the lag operator. If we have more than one  $x$ , all have same lag structure. We can determine length of lag  $p$  by adding lags as long as  $\beta_p$  is statistically significant, and we can choose to max the Akaike information criterion or the Bayesian (Schwartz) information criterion[51].

We can generalize the autoregressive DL models by adding lagged values of  $x$  in addition to lagged  $y$  variables which leads to the **rational lag** model(ADL( $p,q$ )). The ADL( $p,q$ ) model can be written as:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \dots + \beta_p y_{t-p} + \delta_0 x_t + \delta_1 x_{t-1} + \dots + \delta_q x_{t-q} + u_t \text{ or}$$

$$\beta(L)y_t = \beta_0 + \delta(L) x_t + u_t, \text{ thus } y_t = \frac{\beta_0}{\beta(L)} + \frac{\delta(L)}{\beta(L)} x_t + \frac{u_t}{\beta(L)}.$$

Note that, Stationarity depends only on  $\beta(L)$ , not on  $\delta(L)$ . We can easily estimate this model by OLS assuming:

- $E(u_t | y_{t-1}, \dots, y_{t-p}, x_t, x_{t-1}, \dots, x_{t-q}) = 0$ .
- $(y_t, x_t)$  has same mean, variance and autocorrelations for all  $t$ .
- $(y_t, x_t)$  and  $(y_{t-s}, x_{t-s})$  become independent as  $s \rightarrow \infty$ .
- All variables have finite, non-zero fourth moments.
- No perfect multicollinearity.

These are general assumptions that apply to most time-series models.

For the ADL( $p,q$ ) model, we can test Granger causality by testing whether a variable's  $\delta(L)$  polynomial is zero, and we need to leave  $x_t$  out of regression here. the  $F$ -test of set of coefficients on all lags of  $x$ , given effects of lagged  $y$  and any other regressors is used, the rejection means  $x$  causes  $y$ . The Granger causality test whether  $x$  helps to predict  $y$  given the path that  $y$  would follow based on its own lags and perhaps on other regressors. Finally, we can apply all above on MIDAS regressions replacing  $x_t$  notations by  $x_t^{(m)}$  for the high-frequency variable[51].

## **2.10 Estimation Methods**

In this section, we will introduce an overview of some methods of estimation that can be used to estimate parameters of MIDAS regressions. These methods are:

### **2.10.1 The Method of Least squares**

The Least Squares (LS) estimation technique chooses the parameters which minimize the sum of the squared error terms  $Q = \sum_{t=1}^T u_t^2$  assuming that we have  $T$  observations for the low-frequency variable. This method can be used with U-MIDAS regression model:  $y_t = \alpha + \sum_{k=0}^{m-1} \gamma_k L^{\frac{k}{m}} x_t^{(m)} + u_t = \boldsymbol{\gamma}' \mathbf{x}_t^{(m)} + u_t$ , where :

$\mathbf{Y} = (\alpha, \gamma_0, \gamma_1, \dots, \gamma_{m-1})'$ ,  $\mathbf{x}_t^{(m)} = (x_t^{(m)}, x_{t-\frac{1}{3}}^{(m)}, \dots, x_{t-1-\frac{1}{m}}^{(m)})'$  are the explanatory high-frequency variable,  $y_t = (y_1, y_2, \dots, y_T)$  is the low-frequency dependent variable and  $u_t$  are white noise terms such that  $\mathbf{u}_t = (u_1, u_2, \dots, u_T)$ , the white noise process  $u_t$  has zero mean and constant variance  $\sigma^2$ [42,43].

It is easy to show that the solution to this minimization problem gives the following  $(m+1)$  equations:

$$Q = \sum_{t=1}^T \left( y_t - \alpha - \gamma_0 x_t^{(m)} - \gamma_1 x_{t-\frac{1}{3}}^{(m)} - \dots - \gamma_{m-1} x_{t-1-\frac{1}{m}}^{(m)} \right)^2 = 0$$

$$\text{Or } Q = \sum_{k=0}^T \left( y_t - \mathbf{Y}' \mathbf{x}_t^{(m)} \right)^2 = 0$$

This minimization can be accomplished by differentiating these equations with respect to the vector  $\mathbf{Y} = (\alpha, \gamma_0, \gamma_1, \dots, \gamma_{m-1})'$ [42,43].

In the notation above, this procedure gives the normal equations which can be expressed by matrices as the following:

$$\left( \sum_{k=0}^{m-1} \mathbf{x}_t^{(m)} \mathbf{x}_t^{(m)'} \right) \hat{\mathbf{Y}} = \sum_{k=0}^{m-1} \mathbf{x}_t^{(m)} y_t$$

The notations can be simplified by defining  $\mathbf{Y}=y_t$ ,  $\mathbf{X}=\mathbf{x}_t^{(m)}$  and  $\mathbf{u} = u_t$ , then the normal equations can be written as:

$$(\mathbf{X}' \mathbf{X}) \hat{\mathbf{Y}} = \mathbf{X}' \mathbf{Y} \text{ and the solution is } \hat{\mathbf{Y}} = (\mathbf{X}' \mathbf{X})^{-1} \mathbf{X}' \mathbf{Y} \text{ [42, 43].}$$

To show that the least squares estimators have desirable statistical properties it is however necessary to make the following assumptions:

Assumption 1: the model is linear in the parameters and is correctly specified.

Assumption 2:  $E(u_t | x_t^{(m)}, x_{t-\frac{1}{3}}^{(m)}, \dots, x_{t-1-\frac{1}{m}}^{(m)}) = E(u_t) = 0$  for all  $t$ .

Assumption 3:  $u_t$  are uncorrelated with the explanatory variables for all  $t$ .

Assumption 4:  $u_t$  have constant variance and are uncorrelated for all  $t$ .

Assumption 5: there are no exact linear relationships among the explanatory variables(no multicollinearity)[41].



Under these assumptions, it can be shown that the least squares estimators have minimum variance among all linear unbiased estimators. The least squares estimators are said to be best linear unbiased estimators[41].

The sampling distribution of an estimator as  $T$  tends to infinity is its asymptotic distribution. An estimator is said to be asymptotically unbiased if the mean of its asymptotic distribution is equal to the true value of the parameter[41].

An estimator is consistent if the probability that the estimator takes on a value close to the true value goes to 1 as the number of observations  $T$  goes to infinity. Under assumptions 1-5, the least squares estimators are consistent[41].

If we add the assumption that  $u_t$  are normally distributed then the least squares estimators are also normally distributed. Statistical inference on the parameters can then be conducted[41].

### **2.10.2 The Method of Maximum Likelihood**

Maximum Likelihood Estimation (MLE) approach involves forming an assumption about the underlying probability distribution function (pdf) that generates the observed data set, and then estimating parameters of the assumed distribution. Although there are many cases, especially in financial applications, where it may seem inappropriate to assume knowledge of the underlying pdf, White (1982) has been shown that for correctly specified moment equations, the maximum likelihood estimator, now interpreted as Quasi Maximum Likelihood Estimator (QMLE), is consistent. In other words, as long as conditional moments (e.g. mean and/or variance equations) are correctly specified, QMLE will produce estimates that converge to their true parameter values as the sample size increases, although less efficiently than if the correct likelihood function had been used[41].

To illustrate the maximum likelihood approach we consider the previously specified U-MIDAS model now written in the vector notation:

$$y_t = \mathbf{x}_t' \boldsymbol{\gamma} + u_t$$

Where  $\boldsymbol{\gamma}$  is an  $[(m+1) \times 1]$  vector of unknown parameters,  $\mathbf{X} = \mathbf{x}_t^{(m)}$  is an  $[(m+1) \times 1]$  vector of explanatory variables and  $u_t$  is a white noise process with zero mean and variance  $\sigma^2$ . The parameters of interest can then be grouped into a  $[(m+2) \times 1]$  vector  $\boldsymbol{\theta} = \begin{pmatrix} \boldsymbol{\gamma} \\ \sigma^2 \end{pmatrix}$ [41].

The maximum likelihood estimation approach typically involves two steps:

- Specification of a probability distribution for  $u_t$  .

- Computation and maximization of the likelihood function.

The joint *pdf* of the observed sample takes the following form:

$$f(y_1, y_2, \dots, y_T | \mathbf{X}, \boldsymbol{\gamma}, \sigma^2) = f(y_1 | \mathbf{X}, \boldsymbol{\gamma}, \sigma^2) \dots f(y_T | \mathbf{X}, \boldsymbol{\gamma}, \sigma^2) = \prod_{t=1}^T f(y_t | \mathbf{X}, \boldsymbol{\gamma}, \sigma^2) \quad [41].$$

In practice,  $u_t$  is generally assumed to be Gaussian white noise. Assuming normality ( $u_t \sim \text{i. i. d. } N(0, \sigma^2)$ ), the conditional likelihood function of  $y_t$  can be written as

$$f(y_1, y_2, \dots, y_T | \mathbf{X}, \boldsymbol{\gamma}, \sigma^2) = \prod_{t=1}^T (2\pi\sigma^2)^{-1/2} \cdot \exp\left(-\frac{(y_t - \mathbf{x}_t' \boldsymbol{\gamma})^2}{2\sigma^2}\right)$$

The maximum likelihood estimate of  $\boldsymbol{\theta}$  is found by maximizing the above function, which is often interpreted as the probability of observing the realized data sample. In practice we often take logs of the likelihood function in order to simplify algebraic manipulations:

$$L(\boldsymbol{\theta}) = -\frac{T}{2} \ln(2\pi\sigma^2) - \sum_{t=1}^T \left( \frac{(y_t - \mathbf{x}_t' \boldsymbol{\gamma})^2}{2\sigma^2} \right)$$

Where  $L(\boldsymbol{\theta})$  is now known as the log-likelihood function [41].

Maximization of the log-likelihood function involves differentiating  $L(\boldsymbol{\theta})$  with respect to the parameters of interest  $\boldsymbol{\theta}$  and setting each of the resulting equations to zero. Although, in some instances, it may be possible to find a closed form solution to the resultant system of equations, e.g. in the case of a linear regression model, in more complicated situations there are no closed form solutions and  $L(\boldsymbol{\theta})$  must be maximized numerically. In either case, the outcome of maximizing the log-likelihood function results in a maximum likelihood estimate vector  $\boldsymbol{\theta}$  based on a Gaussian likelihood function [41].

In time series analysis, a distinction is made between conditional and exact likelihood functions. The difference is based on the treatment of the first  $p$  observations, where  $p$  is the number of dependent variable lags specified in the model (e.g. in an autoregressive model of order two, AR(2),  $p = 2$ ). The conditional likelihood function assumes that the first  $p$  lags are fixed, that is the likelihood function of the remaining sample is specified conditional on the first  $p$  observations, which are in turn set to either their realized or expected values. The exact likelihood function, on the other hand, is a product of probability densities of all observations, including the first  $p$  lags. Since it is impossible to model the first  $p$  observations using a conditional model (i.e. there is no data to condition the first  $p$  observations in an

AR( $p$ ) model), an unconditional *pdf* is specified for the initial  $p$  observations. When the sample size  $T$  is large, the difference between the two approaches is small[41, 42, 43].

### **2.10.3 Generalized Method Of Moments Estimation**

Generalized Method of Moments (GMM) refers to a class of estimators which are constructed from exploiting the sample moment counterparts of population moment conditions (sometimes known as orthogonality conditions) of the data generating model. GMM estimators have become widely used, for the following reasons:

- GMM estimators have large sample properties that are easy to characterize in ways that facilitate comparison. A family of such estimators can be studied a priori in ways that make asymptotic efficiency comparisons easy. The method also provides a natural way to construct tests which take account of both sampling and estimation error.
- In practice, researchers find it useful that GMM estimators can be constructed without specifying the full data generating process (which would be required to write down the maximum likelihood estimator). This characteristic has been exploited in analyzing partially specified economic models, in studying potentially misspecified dynamic models designed to match target moments, and in constructing stochastic discount factor models that link asset pricing to sources of macroeconomic risk[45].

As we will see, formally there are two alternative ways to specify GMM estimators, but they have a common starting point. Data are a finite number of realizations of the process  $\{x_t : t = 1, 2, \dots\}$ . The model is specified as a vector of moment conditions:

$$g(\gamma_0) = E[f(x_t, \gamma_0)] = \mathbf{0}$$

Where  $\gamma_0$  is a  $K \times 1$  vector of parameters;  $f$  is an  $R$  dimensional vector of functions. If we knew the expectation then we could solve the equations  $g(\gamma_0) = 0$  to find  $\gamma_0$ . If there is a unique solution, so that  $g(\gamma_0) = E[f(x_t, \gamma_0)] = 0$  if and only if  $\gamma = \gamma_0$ . The parameter  $\gamma_0$  is typically not sufficient to write down a likelihood function. Other parameters are needed to specify fully the probability model that underlies the data generation. In other words, the model is only partially specified[44].

For given sample  $x_t : t = 1, 2, \dots, T$ , we cannot calculate the expectation, we replace with sample averages to obtain the analogous sample moments:

$$g_T(\gamma) = \frac{1}{T} \sum_{t=1}^T f(x_t, \gamma)$$

We can derive an estimator  $\hat{\gamma}_{MM}$  as the solution to  $g_T(\hat{\gamma}_{MM}) = 0$ . To find an estimator, we need at least as many equations as we have parameters. For example, in OLS as a MM estimator, we Consider the linear regression model  $y_t$  of on  $x_t$  ( $K \times 1$ ) with white noise error term  $\varepsilon_t$ :

$$y_t = x_t' \boldsymbol{\gamma}_0 + u_t$$

The conditional expectation  $E[y_t | x_t] = x_t' \boldsymbol{\gamma}_0$ , so that  $E[u_t | x_t] = 0$ . That implies the  $K$  unconditional moment conditions  $\boldsymbol{g}(\boldsymbol{\gamma}_0) = E[x_t u_t] = E[x_t (y_t - x_t' \boldsymbol{\gamma}_0)] = 0$ , which we recognize as the minimal assumption for consistency of the OLS estimator. We define the corresponding sample moment conditions as:

$$\begin{aligned} g_T(\hat{\boldsymbol{\gamma}}) &= \frac{1}{T} \sum_{t=1}^T x_t (y_t - x_t' \hat{\boldsymbol{\gamma}}) = \\ &= \frac{1}{T} \sum_{t=1}^T x_t y_t - \frac{1}{T} \sum_{t=1}^T x_t x_t' \hat{\boldsymbol{\gamma}} = 0 \end{aligned}$$

And the MM estimator is derived as the unique solution:

$$\hat{\boldsymbol{\gamma}}_{MM} = \left( \sum_{t=1}^T x_t x_t' \right)^{-1} \sum_{t=1}^T x_t y_t$$

Provided that  $\sum_{t=1}^T x_t x_t'$  is non-singular. Method of moments is one way to motivate the OLS estimator, it highlights the minimal (or identifying) assumptions for OLS[44,45,46].

#### **2.10.4 The Method of Nonlinear Least squares(Gauss-Newton method)**

A straightforward extension of linear specifications is to consider specifications that are nonlinear in parameters. Many empirical evidences show that many economic relationships are in fact nonlinear[49].

Formally, we consider the nonlinear specification:

$$y = f(\mathbf{x}; \boldsymbol{\beta}) + e(\boldsymbol{\beta})$$

Where  $f$  is a given function with  $\mathbf{x}$  an  $m \times 1$  vector of explanatory variables and  $\boldsymbol{\beta}$  is a  $k \times 1$  vector of parameters, and  $e(\boldsymbol{\beta})$  denotes the error of the specification. Note that for a nonlinear specification, the number of explanatory variables  $m$  need not be the same as the number of parameters  $k$ . Given  $T$  observations of  $y$  and  $\mathbf{x}$ , let

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_T \end{bmatrix}, \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\beta}) = \begin{bmatrix} f(\mathbf{x}_1; \boldsymbol{\beta}) \\ f(\mathbf{x}_2; \boldsymbol{\beta}) \\ \vdots \\ f(\mathbf{x}_T; \boldsymbol{\beta}) \end{bmatrix}$$

The nonlinear specification now can be expressed as

$$\mathbf{y} = \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\beta}) + \mathbf{e}(\boldsymbol{\beta}), \text{ where } \mathbf{e}(\boldsymbol{\beta}) \text{ is the vector of errors.}$$

Our objective is to find a  $k$ -dimensional surface that “best” fits the data  $(y_t, \mathbf{x}_t)$ ,  $t = 1, \dots, T$ . Analogous to the OLS method, the method of nonlinear least squares (NLS) suggests to minimize the following NLS criterion function with respect to  $\boldsymbol{\beta}$ :

$$Q_T(\boldsymbol{\beta}) = \frac{1}{T} [\mathbf{y} - \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\beta})]' [\mathbf{y} - \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\beta})] = \frac{1}{T} \sum_t^T [y_t - f(\mathbf{x}_t; \boldsymbol{\beta})]^2$$

The first order condition of the NLS minimization problem is a system of  $k$  nonlinear equations with  $k$  unknowns:

$$\nabla_{\boldsymbol{\beta}} Q_T(\boldsymbol{\beta}) = -\frac{2}{T} \nabla_{\boldsymbol{\beta}} \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\beta}) [\mathbf{y} - \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\beta})] = 0$$

Where

$$\nabla_{\boldsymbol{\beta}} \mathbf{f}(\mathbf{x}_1, \dots, \mathbf{x}_T; \boldsymbol{\beta}) = [\nabla_{\boldsymbol{\beta}} f(\mathbf{x}_1; \boldsymbol{\beta}) \quad \nabla_{\boldsymbol{\beta}} f(\mathbf{x}_2; \boldsymbol{\beta}) \quad \dots \quad \nabla_{\boldsymbol{\beta}} f(\mathbf{x}_T; \boldsymbol{\beta})]$$

Is a  $k \times T$  matrix. When a solution to the NLS minimization problem cannot be obtained analytically, the NLS estimates must be computed using numerical methods. To optimizing a nonlinear function, an iterative algorithm starts from some initial value of the argument in that function and then repeatedly calculates next available value according to a particular rule until an optimum is reached approximately. It should be noted that when there are multiple optima, an iterative algorithm may not be able to locate the global optimum. In fact, it is more common that an algorithm gets stuck at a local optimum, except in some special cases, e.g., when optimizing a globally concave (convex) function. In the literature, several new methods, such as the simulated annealing algorithm, have been proposed to find the global solution. These methods have not yet been standard because they are typically difficult to

implement and computationally very intensive. We will therefore confine ourselves to those commonly used methods[49].

To minimize  $Q_T(\boldsymbol{\beta})$ , a generic algorithm can be expressed as

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} + s^{(i)}\mathbf{d}^{(i)}$$

so that the  $(i+1)^{\text{th}}$  iterated value  $\boldsymbol{\beta}^{(i+1)}$  is obtained from  $\boldsymbol{\beta}^{(i)}$ , the value from the previous iteration, by adjusting the amount  $s^{(i)}\mathbf{d}^{(i)}$ , where  $\mathbf{d}^{(i)}$  characterizes the direction of change in the parameter space and  $s^{(i)}$  controls the amount of change. Different algorithms are resulted from different choices of  $(s)$  and  $(\mathbf{d})$ . As maximizing  $Q_T$  is equivalent to minimizing  $-Q_T$ , the method discussed here is readily modified to the algorithm for maximization problems[49].

Consider the first-order Taylor expansion of  $Q(\boldsymbol{\beta})$  about  $\boldsymbol{\beta}^\dagger$ :

$$Q_T(\boldsymbol{\beta}) \approx Q_T(\boldsymbol{\beta}^\dagger) + [\nabla_{\boldsymbol{\beta}} Q_T(\boldsymbol{\beta}^\dagger)]' (\boldsymbol{\beta} - \boldsymbol{\beta}^\dagger)$$

Replacing  $\boldsymbol{\beta}$  with  $\boldsymbol{\beta}^{(i+1)}$  and  $\boldsymbol{\beta}^\dagger$  with  $\boldsymbol{\beta}^{(i)}$  we have

$$Q_T(\boldsymbol{\beta}^{(i+1)}) \approx Q_T(\boldsymbol{\beta}^{(i)}) + [\nabla_{\boldsymbol{\beta}} Q_T(\boldsymbol{\beta}^{(i)})]' s^{(i)}\mathbf{d}^{(i)}$$

Note that this approximation is valid when  $\boldsymbol{\beta}^{(i+1)}$  is in the neighborhood of  $\boldsymbol{\beta}^{(i)}$ . Let  $\mathbf{g}(\boldsymbol{\beta})$  denote the gradient vector of  $Q_T$ :  $\nabla_{\boldsymbol{\beta}} Q_T(\boldsymbol{\beta})$ , and  $\mathbf{g}^{(i)}$  denote  $\mathbf{g}(\boldsymbol{\beta})$  evaluated at  $\boldsymbol{\beta}^{(i)}$ . If  $\mathbf{d}^{(i)} = -\mathbf{g}^{(i)}$ ,

$$Q_T(\boldsymbol{\beta}^{(i+1)}) \approx Q_T(\boldsymbol{\beta}^{(i)}) - s^{(i)}[\mathbf{g}^{(i)'} \mathbf{g}^{(i)}]$$

As  $\mathbf{g}^{(i)'} \mathbf{g}^{(i)}$  is non-negative, we can find a positive and small enough  $(s)$  such that  $Q_T$  is decreasing. Clearly, when  $\boldsymbol{\beta}^{(i)}$  is already a minimum of  $Q_T$ ,  $\mathbf{g}^{(i)}$  is zero so that no further adjustment is possible. This suggests the following algorithm:

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} - s^{(i)}\mathbf{g}^{(i)}$$

Choosing  $\mathbf{d}^{(i)} = s^{(i)}$  leads to :

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} + s^{(i)}\mathbf{g}^{(i)}$$

which can be used to search for a maximum of  $Q_T$ . The Gauss-Newton method implies to choose the term  $s^{(i)}\mathbf{g}^{(i)} = [\mathbf{D}(\boldsymbol{\beta}^{(i)})' \mathbf{D}(\boldsymbol{\beta}^{(i)})]^{-1} \mathbf{D}(\boldsymbol{\beta}^{(i)})[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta}^{(i)})]$ , where  $\mathbf{D}(\boldsymbol{\beta}) = \nabla_{\boldsymbol{\beta}} \mathbf{f}(\boldsymbol{\beta})$ , so The resulting algorithm is

$$\boldsymbol{\beta}^{(i+1)} = \boldsymbol{\beta}^{(i)} + [\mathbf{D}(\boldsymbol{\beta}^{(i)})' \mathbf{D}(\boldsymbol{\beta}^{(i)})]^{-1} \mathbf{D}(\boldsymbol{\beta}^{(i)})[\mathbf{y} - \mathbf{f}(\boldsymbol{\beta}^{(i)})]$$

Observe that the adjustment term can be obtained as the OLS estimator of regressing  $\mathbf{y} - \mathbf{f}(\boldsymbol{\beta}^{(i)})$  on  $(\boldsymbol{\beta}^{(i)})$ ; this regression is thus known as the Gauss-Newton regression. The iterated  $\boldsymbol{\beta}$  values can be easily computed by performing the Gauss-Newton regression repeatedly[49].

## 2.11 Forecasting

One of the primary objectives of building a model for a time series is to be able to forecast the values for that series at future times. Of equal importance is the assessment of the precision of those forecasts. In this section, we shall consider the calculation of forecasts for MIDAS models. For the most part, we shall assume that the model is known exactly, including specific values for all the parameters. Although this is never true in practice, the use of estimated parameters for large sample sizes does not seriously affect the results[43].

In univariate time series models, the purpose is to forecast the value of  $y_{t+l}$  that will occur  $l$  time units into the future based on the available history of the series up to time  $t$ , namely  $y_1, y_2, \dots, y_t$ . In our work with MIDAS regressions, we have the same purpose with some differences. We want to forecast the future value ( $y_{t+l}$ ) of the low-frequency variable ( $y_t$ ) based on its past values with the existence of the high-frequency variables of other time series related to it. We call time  $t$  the **forecast origin** and  $l$  the **lead time** for the forecast, and denote the forecast itself as  $\hat{y}_t(l)$ [43].

Consider the deterministic trend model

$y_t = \mu_t + u_t$  , where  $\mu_t = \beta_0 + \beta_1 t$  or  $\mu_t = \beta_0 + \beta_1 x_t$ , the stochastic component,  $u_t$ , assumed as white noise with zero mean and constant variance. For this model, we have

$$\begin{aligned}\hat{y}_t(l) &= \hat{y}_{t+l} = E(\hat{\mu}_{t+l} + u_{t+l} | y_1, y_2, \dots, y_t) \\ &= E(\hat{\mu}_{t+l} | y_1, y_2, \dots, y_t) + E(u_{t+l} | y_1, y_2, \dots, y_t) = \hat{\mu}_{t+l} + E(u_{t+l}) = \hat{\mu}_{t+l}\end{aligned}$$

Since for  $l \geq 1$ ,  $u_{t+l}$  is independent of  $y_1, y_2, \dots, y_t$  and has expected value zero. Thus, in this simple case, forecasting amounts to extrapolating the deterministic time trend into the future[43].

Now, Consider the following ADL( $p, m$ )-MIDAS model:

$$y_t = \beta_0 + \sum_{i=1}^p \beta_i L^i y_t + \sum_{k=0}^{m-1} \gamma_k L^{\frac{k}{m}} x_t^{(m)} + u_t = \boldsymbol{\beta}' \mathbf{y} + \boldsymbol{\gamma}' \mathbf{x}^{(m)} + \mathbf{u}$$

where :

$\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_p)'$ ,  $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \dots, \gamma_{m-1})'$ ,  $\mathbf{x}^{(m)} = (x_t^{(m)}, x_{t-\frac{1}{3}}^{(m)}, \dots, x_{t-1-\frac{1}{m}}^{(m)})'$  is the vector of explanatory high-frequency variables,  $\mathbf{y} = (y_{t-1}, y_{t-2}, \dots, y_{t-p})$  is the vector of  $y_t$  lags ( $y_t$  is the low-frequency dependent variable with  $T$  sample size) and  $\mathbf{u} = \mathbf{u}_t = (u_1, u_2, \dots, u_T)$  is the white noise process with zero mean and constant variance  $\sigma^2$ [50].

For this model, we have:

$$\begin{aligned}\hat{y}_t(l) &= \hat{y}_{t+l} = E(y_{t+l} | \mathbf{y}, \mathbf{x}^{(m)}) \\ &= E\left(\hat{\beta}_0 + \sum_{i=1}^p \hat{\beta}_i L^i y_{t+l} + \sum_{k=0}^{m-1} \hat{\gamma}_k L^{\frac{k}{m}} x_{t+l}^{(m)} \mid \mathbf{y}, \mathbf{x}^{(m)}\right) + E(u_{t+l} | \mathbf{y}, \mathbf{x}^{(m)}) \\ &= \hat{\beta}_0 + \sum_{i=1}^p \hat{\beta}_i L^i y_{t+l} + \sum_{k=0}^{m-1} \hat{\gamma}_k L^{\frac{k}{m}} x_{t+l}^{(m)}\end{aligned}$$

Since for  $l \geq 1$ ,  $u_{t+l}$  is independent of  $\mathbf{y}$  and  $\mathbf{x}^{(m)}$ , and has expected value zero. One measurement of forecast accuracy is the Mean Square Forecast Error (MSFE). For simplicity, we consider the following simple MIDAS model:

$$y_t = \mathbf{x}_t^{(m)'} \boldsymbol{\gamma} + \mathbf{u}_t$$

Where  $\mathbf{u}_t$  is white noise process with zero mean and constant variance  $\sigma^2$ [50]. Consider an out of-sample observation  $(y_{t+1}, x_{t+1}^{(m)})$  where  $x_{t+1}^{(m)}$  will be observed but not  $y_{t+1}$ . Given the coefficient estimate  $\hat{\boldsymbol{\gamma}}$ , the standard point estimate of  $E(y_{t+1} | x_{t+1}^{(m)}) = x_{t+1}^{(m)'} \boldsymbol{\gamma}$  is  $\hat{y}_{t+1} = x_{t+1}^{(m)'} \hat{\boldsymbol{\gamma}}$ . The forecast error is the difference between the actual value  $y_{t+1}$  and the point forecast,  $e_{t+1} = y_{t+1} - \hat{y}_{t+1}$ . The mean-squared forecast error (MSFE) is:  $MSFE = E e_{t+1}^2$ . A simple estimator for the MSFE is obtained by averaging the squared prediction errors ( $e_t = y_t - \hat{y}_t$ ):

$$\hat{\sigma}^2 = \frac{1}{T} \sum_{t=1}^T e_t^2$$

The other important forecast accuracy measurement is the Mean Absolute Percentage Error (MAPE) which is defined as:

$$MAPE = \frac{1}{T} \sum_{t=1}^T \frac{|e_t|}{y_t} \times 100\%$$



## **Chapter 3 : Application Case Study**

### **3.1 Modeling The Palestinian GDP**

In Palestine Monetary Authority(PMA) studies, there are many working papers that concentrate on building models to analyze the effect of all kinds of shocks on the economy. On policy analysis, they always need to provide forecasts for many macroeconomic variables in a way that is consistent with the real GDP forecasts. Such models are useful to recite forecasts, to provide a coherent framework, to explain recent developments, to explain the economic outlook on medium term and to explain the main transmission mechanisms of external and internal shocks[35,36,37,38].

The main problem facing these models is that the PMA's structural and financial models are built on annual frequency, because most of the data used in those models are not available on quarterly frequency or it is available in mixed frequencies. Concerning real GDP growth, forecasting the annual models may miss at least part of the influence of the short term economic dynamics, and their short term forecasts may therefore be sub-reliable[35]. Sometimes the data used in PMA studies are available quarterly like the real GDP but its expected factors like the cost of imports (CIM) or the consumer price index(CPI) are available monthly. This makes researchers forced to aggregate monthly data in order to match them with the quarterly GDP, and this may miss part of the influence as mentioned before[36,37].

We will use the idea of MIDAS regressions to present a solution of the mentioned problem, in order to enable building economic and financial models regardless of the availability of data and variables in mixed frequencies.

Due to PMA studies, the Palestinian real GDP can be modeled as a function of exogenous variables(Endogenous variables: the variables that the model seeks to explain or predict from the solution of the model. Exogenous variables: the variables that are determined outside the model). Some of these variables that we are interested in our case study are: The cost of imports(CIM), Employment rate in Israel and settlements(EIS) and Consumer Price Index(CPI) [35,36,37,38]. In their studies, PMA researchers have followed two approaches to build a real Palestinian GDP models based on some indicators. These indicators are expected to have significant impact on economic growth in Palestine. Real GDP models for Palestine can be used to generate essential forecasts. PMA researchers usually use Time-Averaging approach to build a quarterly Palestinian real GDP models based on the main exogenous indicators mentioned above. Only monthly and quarterly variables are

used here, therefore, the monthly indicators are aggregated to be quarterly. Due to PMA, the Palestinian GDP assumed as a small open economy; which means that the external environment is the natural supplier of such exogenous variables on which the local policy makers or economic agents have no control. On the other hand, the external environment in Palestine is very much conditioned by political issues, as well as, economic developments and conditions in Israel. Because that, PMA researchers followed the second approach of building reduced form model that represents a long-run equation for real GDP in Palestine as function of real GDP in Israel and the number of Palestinian residents working in Israel and settlements. The ARDL(Auto Regressive Distributed Lag) technique is used by researchers to build the long-run relationship and to estimate the equilibrium correction model(ECM ) by OLS[35,36].

The inflation model presented in the PMA's 2013 inflation report shows significant long-run relationship between CPI in Palestine and the cost of imports in Palestine(CIM). This result was confirmed by other PMA working paper by Khalil & Dombrecht[36,38].

Based on all the above, we choose to build a quarterly real GDP model in Palestine based on the monthly consumer price index in Palestine(CPI), the monthly cost of imports in Palestine(CIM) and quarterly employment rate in Israel and settlements(EII). In addition, we may include the first lag of GDP( $GDP_{t-1}$ ) and the trend term in the model( $t$ ). We obtained data from the Palestinian Central Bureau of Statistics(PCBS) for a period from 1999 to 2012. The long-run relationship and the estimated short-run ECM regression model can be expressed by the three types of regressions being studied (R-MIDAS regression model, U-MIDAD regression model and Time-Averaging regression model), as the following:

The long-run relationship based on R-MIDAS approach:

$$LGDP_t^{(Q)} = \beta_0 + \beta_1 t + \beta_2 LGDP_{t-1}^{(Q)} + \beta_3 LEII_t^{(Q)} + \beta_4 \sum_{k=0}^{m-1} B_{k,1}(\theta_1, \theta_2) LCPI_{t-\frac{k}{m}}^{(m)} + \beta_5 \sum_{i=0}^{m-1} B_{i,2}(\theta_3, \theta_4) LCIM_{t-\frac{i}{m}}^{(m)} + u_t \quad (\mathbf{R-MIDAS}) \quad (15)$$

Where  $u_t$  is the white noise process(independent identically distributed random variables with zero mean and variance  $\sigma^2$ ),  $t$  is the trend variable,  $B_{k,1}(\theta_1, \theta_2)$  and  $B_{i,2}(\theta_3, \theta_4)$  are lag polynomials that determines the weights for the monthly variables. We take the log for monthly and quarterly variables in the model to

simplify the interpretations of coefficients as percentage change. The estimated short-run ECM regression model based on R-MIDAS approach:

$$\begin{aligned} \widehat{\Delta LGDP}_t^{(Q)} = & \hat{c} + \hat{\gamma} \hat{u}_{t-1} + \hat{\alpha}_1 \widehat{LGDP}_{t-1}^{(Q)} + \hat{\alpha}_2 \Delta^{(Q)} LEII_t^{(Q)} + \\ & \hat{B}_1 \sum_{k=0}^{m-1} B_{k,3}(\hat{\theta}_5, \hat{\theta}_6) \Delta^{(m)} LCPI_{t-\frac{k}{m}}^{(m)} + \\ & \hat{B}_2 \sum_{i=0}^{m-1} B_{i,4}(\hat{\theta}_7, \hat{\theta}_8) \Delta^{(m)} LCIM_{t-\frac{i}{m}}^{(m)} \end{aligned} \quad (16)$$

Where  $\hat{u}_{t-1}$  is the first lag of the estimated residuals of the long-run model and its coefficient  $\hat{\gamma}$  denotes to the equilibrium or adjustment rate(see sections 2.7 and 2.8 of cointegration). The symbol  $\Delta^{(m)}$  means that we are taking the first difference of the monthly variable or the high-frequency variable, and the symbol  $\Delta^{(Q)}$  means that we are taking the first difference of the quarterly variable or the low-frequency variable.

The long-run relationship based on U-MIDAS approach:

$$\begin{aligned} LGDP_t^{(Q)} = & \beta_0 + \beta_1 t + \beta_2 LGDP_{t-1}^{(Q)} + \beta_3 LEII_t^{(Q)} + \sum_{k=0}^{m-1} \beta_{k,4} LCPI_{t-\frac{k}{m}}^{(m)} + \\ & \sum_{i=0}^{m-1} \beta_{i,5} LCIM_{t-\frac{i}{m}}^{(m)} + u_t \end{aligned} \quad (\text{U-MIDAS}) \quad (17)$$

Note that  $\beta_{k,4}$  and  $\beta_{i,5}$  here are different coefficients for all  $k$  and  $i$ , and they are not lag polynomials as in R-MIDAS model. The estimated short-run ECM regression model based on U-MIDAS approach:

$$\begin{aligned} \widehat{\Delta LGDP}_t^{(Q)} = & \hat{c} + \hat{\gamma} \hat{u}_{t-1} + \hat{\alpha}_1 \widehat{LGDP}_{t-1}^{(Q)} + \hat{\alpha}_2 \Delta^{(Q)} LEII_t^{(Q)} + \\ & \sum_{k=0}^{m-1} \hat{B}_{k,1} \Delta^{(m)} LCPI_{t-\frac{k}{m}}^{(m)} + \sum_{i=0}^{m-1} \hat{B}_{i,2} \Delta^{(m)} LCIM_{t-\frac{i}{m}}^{(m)} \end{aligned} \quad (18)$$

Also here  $\hat{B}_{k,1}$  and  $\hat{B}_{i,2}$  are different for all  $k$  and  $i$ , and they are not estimated as lag polynomials.

The long-run relationship based on Time-Averaging approach:

$$\begin{aligned} LGDP_t^{(Q)} = & \beta_0 + \beta_1 t + \beta_2 LGDP_{t-1}^{(Q)} + \beta_3 LEII_t^{(Q)} + \beta_4 LCPI_t^{(Q)} + \\ & \beta_5 LCIM_t^{(Q)} + u_t \end{aligned} \quad (\text{Time-Averaging}) \quad (19)$$

Where:

$$LCPI_t^{(Q)} = \log\left\{\frac{1}{m} \sum_{k=0}^{m-1} CPI_{t-\frac{k}{m}}^{(m)}\right\} \text{ and } LCIM_t^{(Q)} = \log\left\{\frac{1}{m} \sum_{i=0}^{m-1} CIM_{t-\frac{i}{m}}^{(m)}\right\}.$$

And the estimated short-run ECM regression model based on Time-Averaging approach:

$$\widehat{\Delta LGDP}_t^{(Q)} = \hat{c} + \hat{\gamma} \hat{u}_{t-1} + \hat{\alpha}_1 \widehat{LGDP}_{t-1}^{(Q)} + \hat{\alpha}_2 \Delta^{(Q)} LEII_t^{(Q)} + \hat{B}_1 \Delta^{(Q)} LCPI_t^{(Q)} + \hat{B}_2 \Delta^{(Q)} LCIM_t^{(Q)} \quad (20)$$

Where:

$$\Delta^{(Q)} LCPI_t^{(Q)} = \Delta^{(Q)} \log\left\{\frac{1}{m} \sum_{k=0}^{m-1} CPI_{t-\frac{k}{m}}^{(m)}\right\} \text{ and}$$

$$\Delta^{(Q)} LCIM_t^{(Q)} = \Delta^{(Q)} \log\left\{\frac{1}{m} \sum_{i=0}^{m-1} CIM_{t-\frac{i}{m}}^{(m)}\right\}$$

Our results of time series analysis are shown in appendix A. The AGK test of all  $\theta$ 's are zeros(see Table A6 in Appendix A) were accepted, the null hypothesis is  $H_0 : \theta = 0$ , ( $\theta=(\theta_1, \theta_2)$ )[32]. In other words, accepting the null hypothesis implies that the estimated R-MIDAS regression does not differ from time-averaging aggregation(see appendix C for more details about the AGK test)[32]. This implies to use time-averaging since all weights will be equal and we conclude that each Exponential Almon Lag polynomial weight equals to  $1/m$ (because  $m=3$  in our analysis,  $B_{i,j}(k; \theta) = \frac{1}{3}$ , for all  $i=1,2,3$  and  $j=1,2$ ). Thus, the results show that both the long-run relationship based on R-MIDAS regression and that based on Time-Averaging regression model do the same work here. The results of estimation of the long-run model based on the Time-Averaging method(see Table A7) exhibited that: the Employment rate in Israel and settlements(EII), the first lag of GDP and the cost of imports in Palestine(CIM) all affect the real Palestinian GDP.

The results of U-MIDAS regression were better(see Table A3). The results exhibited that the effect of the cost of imports (CIM) on the real GDP comes from the second-month based series of the imports( $LCIM_{t-\frac{1}{3}}^{(3)}$  or (LCIM1) because the others( $LCIM_t^{(3)}$ ,  $LCIM_{t-\frac{2}{3}}^{(3)}$ ) or (LCIM0, LCIM2) are not significant. We found that in both Time-Averaging and U-MIDAS methods, the monthly CPI (the first, the second and the third month based-series) have no effect on GDP(see Table A3 and Table A7), so we decided to eliminate the monthly CPI from the long-run regression, and we decided to eliminate both the first-month based series of imports and the third-

month based series of imports( $LCIM_t^{(3)}$  and  $LCIM_{t-\frac{2}{3}}^{(3)}$ ) from the model. The test of stationarity of all variables performed by Augmented Dickey-Fuller test(see Table A1). The results demonstrated that variables are integrated of order one, and the residuals were stationary due to U-MIDAS and Time-Averaging reduced models(after eliminating CPI)(see Table A9), so the test of cointegration by Engle-Granger procedures was satisfied(see section 2.7 for cointegration tests). The test of Johansen suggests that there exist at least one cointegration relationship and at most two(see table A2). So we concluded that there exist long-run relationship between the GDP and all of the other variables except the CPI.

Based on the above results, we estimated the equilibrium correction model (ECM) in order to construct the short-run relationship using the three methods: U-MIDAS, R-MIDAS and Time-Averaging. As expected, both R-MIDAS and Time-Averaging will do the same work in estimation of the ECM model because the null hypothesis of the AGK test was accepted(see Table A10). Also we found that each ECM based on each method exhibited significant dependency between the differenced GDP and the differenced cost of imports only(see Tables A10, A11 and A13). We noted from the table A11 in Appendix A that the estimated ECM based on U-MIDAS approach outperformed the estimated ECM based on Time-Averaging approach; the U-MIDAS approach showed significant dependency between the first differenced GDP and the second differenced monthly cost of imports( $\Delta^{(3)}LCIM_{t-\frac{2}{3}}^{(3)}$ ), but we expect that there is also significant dependency between the first differenced GDP and the first differenced monthly cost of imports( $\Delta^{(3)}LCIM_{t-\frac{1}{3}}^{(3)}$ ), so we will include this term in the final ECM. Finally, we estimated the reduced ECM after eliminating all non-significant terms(except  $\Delta^{(3)}LCIM_{t-\frac{1}{3}}^{(3)}$  and  $\Delta^{(3)}LCIM_{t-\frac{2}{3}}^{(3)}$ ) based on U-MIDAS approach as the following(see Table A12):

$$\widehat{\Delta LGDP}_t^{(Q)} = (0.0022) + (-0.399) u_{t-1} + (0.194) \Delta^{(3)}LCIM_{t-\frac{1}{3}}^{(3)} + (0.104) \Delta^{(3)}LCIM_{t-\frac{2}{3}}^{(3)} \quad (21)$$

Note that we eliminated the differenced third-month based series of the cost of imports( $\Delta^{(3)}LCIM_t^{(3)}$ ) from the estimated ECM, because it has no impact on the differenced GDP(not significant). Also the following variables were eliminated because the same reason: The first lag of quarterly GDP ( $LGDP_{t-1}^{(Q)}$ ), the quarterly

differenced employment rate in Israel and settlements ( $\Delta^{(Q)}LEII_t^{(Q)}$ ), and the differenced monthly consumer price index with all its parts ( $\widehat{B}_{i,1} \Delta^{(m)}LCPI_{t-\frac{k}{3}}^{(m)}$ ,  $k=0,1,2$ ).

From the estimated ECM equation (21) and by the results of Table A12, it is clear that the coefficient of error correction term ( $u_{t-1}$ ) has negative sign and it is statistically significant ( $p$ -value  $< 0.05$ ). Based on our results, we conclude that there is movement back of the real Palestinian GDP to the equilibrium relationship. This means that if the real Palestinian GDP exceeds the long-run equilibrium relationship, then the error-correction term will help to move the real GDP back to the original equilibrium. Also if real GDP is lower than the long-run equilibrium level, then the error-correction term will help to shift the real GDP toward the long-run equilibrium relationship.

The estimated value of  $u_{t-1}$  coefficient ( $\hat{\gamma} = -0.399158$ ) implies that about 39.9% of the deviation between real GDP and the long-run equilibrium value is reduced every quarter, and the adjustment will take about two quarters and half ( $100/39.9 = 2.51$ ). With respect to the impact on short-run, coefficient values of the differenced second-month based series and the differenced first-month based series of the cost of imports ( $\Delta^{(3)}LCIM_{t-\frac{1}{3}}^{(3)}$  and  $\Delta^{(3)}LCIM_{t-\frac{2}{3}}^{(3)}$ ) provide the evidence that there is short-run effect of these monthly-based series on real Palestinian GDP.

On the other hand, the differenced first-month based series has not effect on the short-run, because its coefficient is not significant based on the results on Table A11, and so it was eliminated from the ECM equation (21). From equation (21), it is concluded that on the short-run, we expect that the real GDP in each quarter will increase by 1.94% on the average when the cost of imports at the second month in that quarter increases by 10% with other factors are held constant. Also, we expect that the real GDP in each quarter will increase by 1.04% on the average when the cost of imports at the first month in that quarter increases by 10% with other factors are held constant.

In summary, we have identified that there exists long-run equilibrium relationship between both Employment rate in Israel and settlements (EII) and the first lag of GDP and the real Palestinian GDP. Also, the real Palestinian GDP is affected by the second-month based series of the cost of imports in Palestine ( $LCIM_{t-\frac{1}{3}}^{(3)}$ ). The estimated cointegrating equation using U-MIDAS regression is (see Table A4):

$$LGDP_t^{(Q)} = (5.639) + (0.004)t + (0.368)LGDP_{t-1}^{(Q)} + (0.144)LEII_t^{(Q)} + (0.219)LCIM_{t-\frac{1}{3}}^{(3)} \quad (22)$$

We note that each variable in the long-run equation (22) causes the real GDP to increase. On the long-run, we expect that each ten percentage increase in the cost of imports at the second month in each quarter will increase the quarterly real GDP by 2.19% on the average with other variables are held constant. Also we expect that each ten percentage increase in the quarterly employment rate in Israel and settlements will increase the quarterly real GDP by 1.44% on the average with other variables are held constant. Similarly, we expect that the real GDP will increase by 3.68% on the average for each 10% increasing of its first lag with other variables are held constant. Finally, it is concluded that the real GDP in Palestine is increasing by only 0.04% on the average through the time with other variables are held constant .\*

Table (3) state out-of-sample (that is, part of model estimation sample) forecasts of GDP for the quarters of the year 2012. The table compares these forecasts with the true observed values. The table shows smaller forecast errors when using MIDAS models(restricted and unrestricted) than when using Time-Averaging model. The absolute percentage errors in the forecasts of real GDP for these out of sample part when using U-MIDAS model is about 4.2% on the average, and it is lower than that by the other methods(for more details about MAPE see section 2.11).

**Table (3): Quarterly out-of-sample forecasts of GDP<sub>t</sub>\***

Year : Quarter	Observed	U-MIDAS		R-MIDAS		Time-Averaging	
		Forecast	Error	Forecast	Error	Forecast	Error
<b>2012:Q1</b>	1.602	1.564	0.038	1.556	0.045	1.548	0.054
<b>2012:Q2</b>	1.755	1.605	0.150	1.603	0.151	1.584	0.171
<b>2012:Q3</b>	1.708	1.707	0.001	1.690	0.018	1.669	0.039
<b>2012:Q4</b>	1.733	1.630	0.103	1.643	0.090	1.623	0.110
<b>MAPE(out-of-sample)</b>		<b>4.2%</b>		<b>4.4%</b>		<b>5.4%</b>	

\* values in \$1000000.

We used the estimated long-run relationship (22) to forecast the value of the Palestinian real GDP at the first quarter in the year 2013(2013:Q1). The forecasted

\* The data used in analysis were collected from Palestinian Central Bureau of Statistics(PCBS).

value was 1425043, comparing with the true value(1644800), we have 13% forecast error for the Palestinian real GDP at the first quarter in the year 2013.

### **3.2 Discussion of Results and Conclusions**

The assumption of equal frequencies for all the variables is not always available or satisfied for the classical regression models. For example, in order to explore the impact of some factors on the Gross Domestic Product-the most important measure of a nation's economic health- or for forecasting issues, we need all variables to be in the same frequencies. Economic variables are usually available with various or mixed frequencies(annually, quarterly, monthly...etc ); thus it is difficult to deal with these variables directly. The classical method is to take the simple average to convert high-frequency variables(more frequently observed than the response) into fixed frequency variables, and to make them identical in frequencies with the low-frequency variable(less frequently observed than factors), which is assumed to be dependent or response. We introduced in this thesis a relatively recent method established by Ghysels, Santa-Clara and Valkanov(2004) that is Mixed Data Sampling (MIDAS) regression model. This method coincides the frequency of the response and the regressors, and also makes the application of common statistical procedures possible. We explained the features of MIDAS regressions which often outperform the typical approach of time aggregation in terms of estimation efficiency(less error variance)[32]. We saw that MIDAS regressions have two main types, unrestricted and restricted. The later characterized by diversity of parameterization methods in order to estimate the model coefficients, and it supports the parsimony principle which is important to simplify interpretations.

The performance of the three approaches(Unrestricted MIDAS regression, Restricted MIDAS regression and the classical Time-Averaging regression) was assessed through modeling the Palestinian GDP with some suggested factors that may affect it significantly. We compared between these approaches by constructing long-run relationships using the three methods. The aim was to know which one is the best for reducing forecasting errors. Our results showed that both U-MIDAS and R-MIDAS were better than the classical Time-Averaging method in reducing forecasting errors measured by the mean absolute percent error(MAPE) as shown from the Table(3), this result was compatible with many previous studies. Also, our results showed that U-MIDAS was better than R-MIDAS narrowly in reducing forecasting errors; because the difference between frequencies is not large(one quarter = three months)[6]. Moreover, we have seen that both U-MIDAS and R-MIDAS approaches separate the high-frequency variable into several variables; Each



variable has the same frequency as the low-frequency response variable, then it assigns group of parameters or weights to them, we estimated the weights of the R-MIDAS models in our case study using Exponential Almon Lag Polynomial by NLS method, also we estimated the coefficients of the U-MIDAS models in our case study using by OLS method. These approaches guarantees reserving the high-frequency information and gives clearer results than the classical Time-Averaging method.

We have noticed all the benefits that mentioned previously by the results of the data analysis. The results showed that the monthly Consumer Price Index(CPI) has no significant impact on GDP, since each monthly series of CPI has no significant impact on GDP, but concerning the monthly Cost of Imports(CIM), the results were different. The CIM has significant impact on GDP; specifically the second-month-based series of CIM, which means that the quarterly GDP in the long-run affected by its cost of imports in the second month computed in each quarter. Also, the results turned out that in the long-run, the quarterly GDP has increasing general trend and it is affecting by its first lag and the quarterly Employment Rate In Israel and Settlements, also, all these variables have positive impact on GDP. In the short-run, after estimating the equilibrium correction model; the results showed that the GDP in the short-run is affecting by the CIM; specifically by the second-month-based series and the first-month-based series of CIM. The speed of adjustment is about 40%; which means that we need about 7 months(2.5 quarter) to shift the real GDP toward the long-run equilibrium relationship with the CIM. All these results about the long-run relationship and the short-run relationship were consistent with the PMA studies[35,36,37,38].

Possibilities for future study include the application of MIDAS approach in forecasting the real GDP of Palestine for the period 1996 to 2000, because the GDP at that period is not available neither quarterly nor annually. Finally, other important topics related to time series analysis such as causality and disaggregation can be further introduced with MIDAS approach; specifically Granger causality and disaggregation using MIDAS regressions[26,39].

## **Appendix A : Output Tables**

**Table A1: Unit root test**

variable	Frequency type	ADF test p-value	ADF test p-value for first difference
GDP	Quarterly	0.05469	0.01667
EII	Quarterly	0.2279	<0.01
CPI0	Monthly	0.2336	0.04454
CPI1	Monthly	0.3104	0.02248
CPI2	Monthly	0.3837	0.0191
CIM0	Monthly	0.482	<0.01
CIM1	Monthly	0.02025	<0.01
CIM2	Monthly	0.5246	<0.01

**Table A2: Results of Johansen's Cointegration Test\***

Trace test	Test statistic	10%	5%	1%	Results
H0: $r \leq 4$ , H1: $r > 4$	4.80	7.52	9.24	12.97	Fail to reject H0
H0: $r \leq 3$ , H1: $r > 3$	14.06	17.85	19.96	24.60	Fail to reject H0
H0: $r \leq 2$ , H1: $r > 2$	25.03	32.00	34.91	41.07	Fail to reject H0
H0: $r \leq 1$ , H1: $r > 1$	57.72	49.65	53.12	60.16	reject H0
H0: $r = 0$ , H1: $r > 0$	108.46	71.86	76.07	84.45	reject H0
Max-eigenvalue test	Test statistic	10%	5%	1%	Results
H0: $r \leq 4$ , H1: $r = 5$	4.80	7.52	9.24	12.97	Fail to reject H0
H0: $r \leq 3$ , H1: $r = 4$	9.26	13.75	15.67	20.20	Fail to reject H0
H0: $r \leq 2$ , H1: $r = 3$	10.97	19.77	22.00	26.81	Fail to reject H0
H0: $r \leq 1$ , H1: $r = 2$	32.69	25.56	28.14	33.24	reject H0
H0: $r = 0$ , H1: $r = 1$	50.74	31.66	34.40	39.79	reject H0

\*  $r$  is the number of cointegrating relationships.

\*variables: (gdp,eii,cim0,cim1,cim2).

**Table A3: Estimation Results for Long-Run U-MIDAS Regression Equation (17)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$\beta_0$	8.159787	2.885871	2.827	<0.01*
Trend	$\beta_T$	0.010546	0.006514	1.619	0.112433
LGDP <sub>t-1</sub>	$\beta_1$	0.362917	0.089189	4.069	<0.01*
LEII	$\beta_2$	0.141811	0.060549	2.342	0.023661*
LCPI0	$\beta_{0,3}$	0.454766	1.429812	0.318	0.751911
LCPI1	$\beta_{1,3}$	-2.59173	2.239979	-1.15	0.253363
LCPI2	$\beta_{2,3}$	1.378244	1.429546	0.964	0.340142
LCIM0	$\beta_{0,4}$	0.092053	0.068231	1.349	0.184046
LCIM1	$\beta_{1,4}$	0.148495	0.061672	2.408	0.02021*
LCIM2	$\beta_{2,4}$	0.0543	0.075718	0.717	0.477
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.05888	0.9356	0.9227	72.61		2.2e-16
df			df1	df2	
45			9	45	

**Note:** This table contains estimation results of the **U-MIDAS** regression equation (17) with trend variable with coefficient  $\beta_T$ . The dependent variable is LGDP.

**Table A4: Estimation Results for Long-Run U-MIDAS Regression Equation (22)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$\beta_0$	5.639495	1.017943	5.540	<0.001*
Trend	$\beta_T$	0.004340	0.001421	3.053	<0.001*
LGDP <sub>t-1</sub>	$\beta_1$	0.368372	0.084775	4.345	<0.001*
LEII	$\beta_2$	0.143993	0.051167	2.814	<0.01*
LCIM1	$\beta_{1,4}$	0.219608	0.046672	4.705	<0.001*
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.05971	0.9264	0.9205	157.3		2.2e-16
df			df1	df2	
50			4	50	

**Note:** This table contains estimation results of the **reduced form** of the **U-MIDAS** regression equation (17) with trend variable with coefficient  $\beta_T$ . The dependent variable is LGDP.

**Table A5: Estimation Results for Long-Run R-MIDAS Regression Equation (15)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$\beta_0$	8.486	2.933	2.893	<0.01*
Trend	$\beta_T$	1.073e-02	6.621e-03	1.620	0.112262
LGDP <sub>t-1</sub>	$\beta_1$	3.346e-01	9.065e-02	3.691	<0.001*
LEII	$\beta_2$	1.541e-01	6.154e-02	2.505	0.015949*
LCPI0	$\beta_3$	-7.681e-01	5.935e-01	-1.294	0.202163
LCPI1	$\theta_1$	-5.146e-01	2.044e+09	0.000	1.00000
LCPI2	$\theta_2$	-2.421	6.814e+08	0.000	1.00000
LCIM0	$\beta_4$	3.015e-01	7.273e-02	4.146	<0.001*
LCIM1	$\theta_3$	1.734	3.676	0.472	0.639378
LCIM2	$\theta_4$	-4.773e-01	9.755e-01	-0.489	0.627044
Residual standard error			df		
0.05984			45		

**Note:** This table contains estimation results of the **R-MIDAS** regression equation (15) with trend variable with coefficient  $\beta_T$ . Started values used in NLS estimation method :  $(\beta_3, \theta_1, \theta_2) = (\beta_4, \theta_3, \theta_4) = (0, 0, 0)$ . The dependent variable is LGDP.

**Table A6: Estimation Results for Long-Run R-MIDAS Regression Equation (15)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$\beta_0$	8.279	3.589	2.307	0.0257*
Trend	$\beta_T$	1.027e-02	7.859e-03	1.307	0.1979
LGDP <sub>t-1</sub>	$\beta_1$	3.383e-01	7.762e-02	4.359	<0.001*
LEII	$\beta_2$	1.509e-01	5.671e-02	2.661	0.0108*
LCPI0	$\beta_3$	-7.154e-01	7.574e-01	-0.945	0.3499
LCPI1	$\theta_1$	-2.376e-01	1.835e+05	0.000	1.000000
LCPI2	$\theta_2$	-1.309	6.112e+04	0.000	1.000000
LCIM0	$\beta_4$	2.958e-01	6.848e-02	4.319	<0.001*
LCIM1	$\theta_3$	2.802	3.095	0.906	0.3700
LCIM2	$\theta_4$	-7.518e-01	8.830e-01	-0.851	0.3990
Residual standard error	df	AGK test statistic/df	p-value	conclusion	
0.05975	45	0.3307/6	0.9993	Accept $H_0 : \theta = 0$	

**Note:** This table contains estimation results of the **R-MIDAS** regression equation (15) with trend variable with coefficient  $\beta_T$ . Started values used in NLS estimation method were taken from the first estimated values:  $(\beta_3, \theta_1, \theta_2) = (-5.995e-01, -3.161e-01, -1.544)$ ,  $(\beta_4, \theta_3, \theta_4) = (3.095e-01, 2.217, -5.910e-01)$ . The dependent variable is LGDP.

**Table A7: Estimation Results for Long-Run Time-Averaging Regression Equation (19)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$\beta_0$	8.120633	2.630812	3.087	<0.01*
Trend	$\beta_T$	0.009919	0.005930	1.673	0.100748
LGDP <sub>t-1</sub>	$\beta_1$	0.333306	0.083688	3.983	<0.001*
LEII	$\beta_2$	0.147799	0.055264	2.674	0.010145*
LCPI	$\beta_3$	-0.711832	0.539911	-1.318	0.193491
LCIM	$\beta_4$	0.314057	0.064613	4.861	<0.001*
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.05749	0.9331	0.9263	136.7		2.2e-16
df			df1	df2	
49			5	49	

**Note:** This table contains estimation results of the **Time-Averaging** regression equation (19) with trend variable with coefficient  $\beta_T$ . The dependent variable is LGDP.

**Table A8: Estimation Results for Long-Run Time-Averaging Regression (reduced form of Equation (19))**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$\beta_0$	4.913330	1.008993	4.870	<0.001*
Trend	$\beta_T$	0.002378	0.001576	1.509	0.137661
LGDP <sub>t-1</sub>	$\beta_1$	0.330878	0.084283	3.926	<0.001*
LEII	$\beta_2$	0.119393	0.051266	2.329	0.023946*
LCIM	$\beta_4$	0.330163	0.063914	5.166	<0.001*
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.05791	0.9307	0.9252	168		2.2e-16
df			df1	df2	
50			4	50	

**Note:** This table contains estimation results of the **reduced form** of the **U-MIDAS** regression equation (19) with trend variable with coefficient  $\beta_T$ . The dependent variable is LGDP.

**Table A9: Unit root test for Error Correction Term  $u_t$** 

Method	ADF test p-value	PP test p-value
U-MIDAS	0.03943	<0.01
Time-Averaging	0.07439	<0.01
R-MIDAS	0.04703	<0.01

**Note:** This table contains **Augmented Dickey–Fuller(ADF)** test and **Phillips-Perron(PP)** test results of the estimated residuals term  $\hat{u}_t$  resulted by the reduced forms of the **U-MIDAS** and **Time-Averaging** regressions.

**Table A10: Estimation Results for Short-Run ECM-RMIDAS Equation (16)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$c$	-0.14864	0.25114	-0.592	0.556968
$u_{t-1}$	$\gamma$	-0.35123	0.12667	-2.773	<0.01*
LGDP <sub>t-1</sub>	$\alpha_1$	0.01187	0.01774	0.669	0.506955
$\Delta^{(Q)}$ LEII	$\alpha_2$	0.04109	0.05712	0.719	0.475792
$\Delta^{(m)}$ LCPI	$B_1$	-1.53490	0.66262	-2.316	0.025255*
	$\theta_5$	0.15474	30.86770	0.005	0.996023
	$\theta_6$	0.18553	6.87614	0.027	0.978596
$\Delta^{(m)}$ LCIM	$B_2$	0.31570	0.08905	3.545	<0.001*
	$\theta_7$	4.66255	3.92171	1.189	0.240853
	$\theta_8$	-1.12753	0.90617	-1.244	0.219983
Residual standard error	df	AGK test statistic/df	p-value	conclusion	
0.05558	44	0.4917/6	0.9979	Accept $H_0 : \theta = 0$	

**Note:** This table contains estimation results of the **ECM-RMIDAS-based** equation (16). Started values used in NLS estimation method:  $(B_1, \theta_5, \theta_6) = (B_2, \theta_7, \theta_8) = (0, 0, 0)$ . The dependent variable is  $\Delta^{(Q)}$ LGDP.

**Table A11: Estimation Results for Short-Run ECM-UMIDAS Equation (18)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$c$	-0.16068	0.53923	-0.298	0.7671
$\mathbf{u}_{t-1}$	$\gamma$	-0.35056	0.14740	-2.378	0.0218*
LGDP <sub>t-1</sub>	$\alpha_1$	0.01271	0.03857	0.329	0.7434
$\Delta^{(Q)}$ LEII	$\alpha_2$	0.04145	0.05256	0.789	0.4346
$\Delta^{(m)}$ LCPI0	$B_{0,1}$	0.10126	0.97738	0.104	0.9180
$\Delta^{(m)}$ LCPI1	$B_{1,1}$	-0.82621	1.46904	-0.562	0.5767
$\Delta^{(m)}$ LCPI2	$B_{2,1}$	-0.76059	1.05557	-0.721	0.4750
$\Delta^{(m)}$ LCIM0	$B_{0,2}$	0.04761	0.06563	0.725	0.4720
$\Delta^{(m)}$ LCIM1	$B_{1,2}$	0.19154	0.04155	4.610	<0.001*
$\Delta^{(m)}$ LCIM2	$B_{2,2}$	0.06785	0.05395	1.258	0.2151
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.0555	0.6146	0.5357	7.795		9.396e-07
df			df1	df2	
44			9	44	

**Note:** This table contains estimation results of the **ECM-UMIDAS-based equation (18)**. The dependent variable is  $\Delta^{(Q)}$ LGDP.

**Table A12: Estimation Results for Short-Run ECM-UMIDAS Equation (21)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$c$	0.002212	0.007489	0.295	0.76895
$\mathbf{u}_{t-1}$	$\gamma$	-0.399158	0.132621	-3.010	<0.01*
$\Delta^{(m)}$ LCIM1	$B_{1,2}$	0.193961	0.040016	4.847	<0.001*
$\Delta^{(m)}$ LCIM2	$B_{2,2}$	0.104397	0.042225	2.472	0.01687*
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.0549	0.5715	0.5458	22.23		2.76e-09
df			df1	df2	
50			3	50	

**Note:** This table contains estimation results of the **reduced form** of the **ECM-UMIDAS-based equation (18)**. The dependent variable is  $\Delta^{(Q)}$ LGDP.

**Table A13: Estimation Results for Short-Run ECM-Time-Averaging Equation (20)**

variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$c$	-0.070002	0.533555	-0.131	0.8962
$\mathbf{u}_{t-1}$	$\gamma$	-0.284815	0.140779	-2.023	0.0486*
LGDP <sub>t-1</sub>	$\alpha_1$	0.006112	0.038163	0.160	0.8734
$\Delta^{(Q)}$ LEII	$\alpha_2$	0.044597	0.042387	1.052	0.2980
$\Delta^{(Q)}$ LCPI	$B_1$	-1.423650	0.806855	-1.764	0.0840
$\Delta^{(Q)}$ LCIM	$B_2$	0.373267	0.058953	6.332	<0.001*
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.05511	0.5855	0.5423	13.56		2.952e-08
df			df1	df2	
48			5	48	

**Note:** This table contains estimation results of the **ECM-Time-Averaging -based** equation (20). The dependent variable is  $\Delta^{(Q)}$ LGDP.

**Table A14: Estimation Results for Short-Run ECM-Time-Averaging (reduced form of Equation (20))**

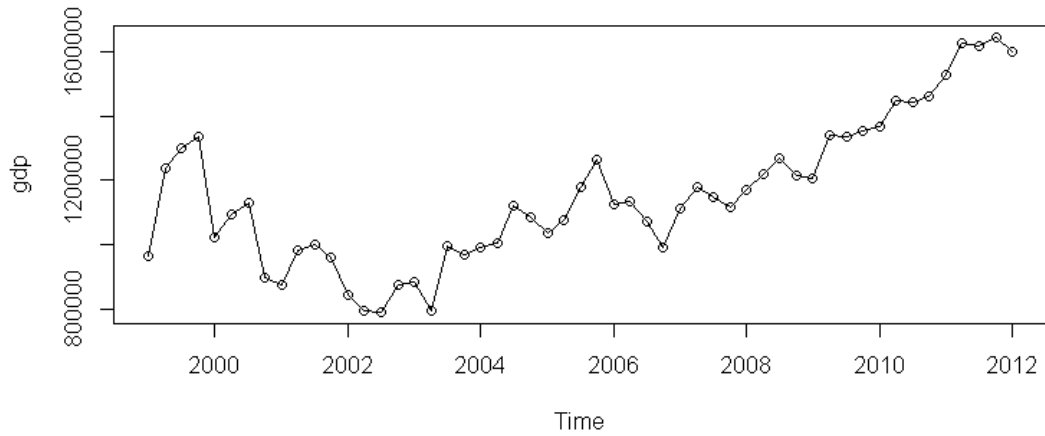
variable	parameter	Estimate	Std. Error	t value	Pr(> t )
Intercept	$c$	0.0009689	0.0076079	0.127	0.899
$\mathbf{u}_{t-1}$	$\gamma$	-0.3348287	0.1329698	-2.518	0.015 *
$\Delta^{(Q)}$ LCIM	$B_2$	0.4125951	0.0539719	7.645	<0.001*
Residual standard error	Multiple R-squared	Adjusted R-squared	F-statistic		p-value
0.05569	0.5502	0.5325	31.19		1.421e-09
df			df1	df2	
51			2	51	

**Note:** This table contains estimation results of the **reduced form** of the **ECM-Time-Averaging -based** equation (20). The dependent variable is  $\Delta^{(Q)}$ LGDP.

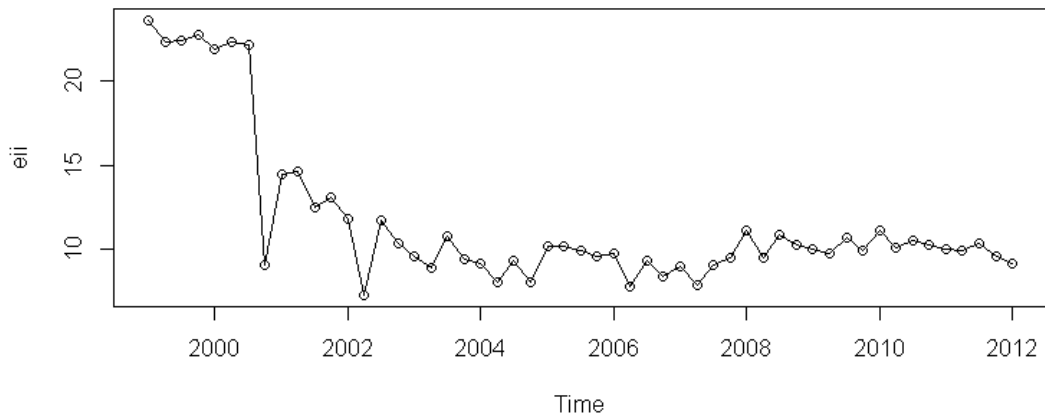


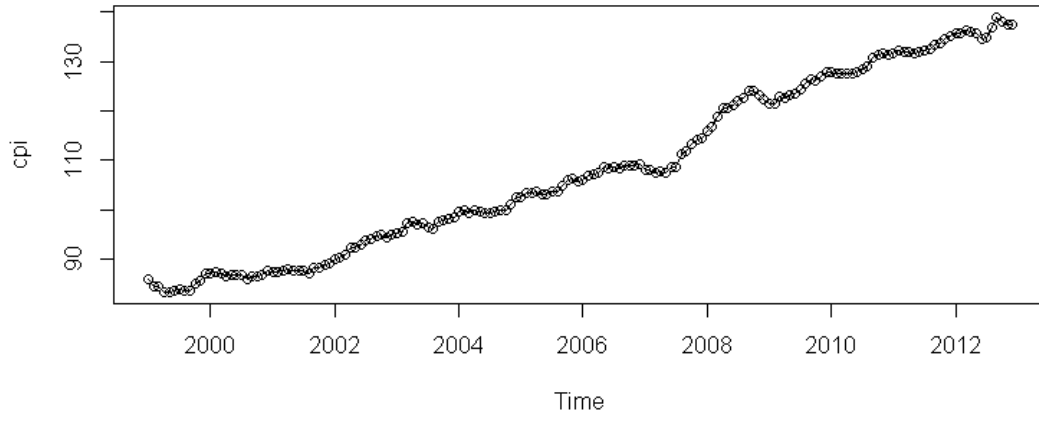
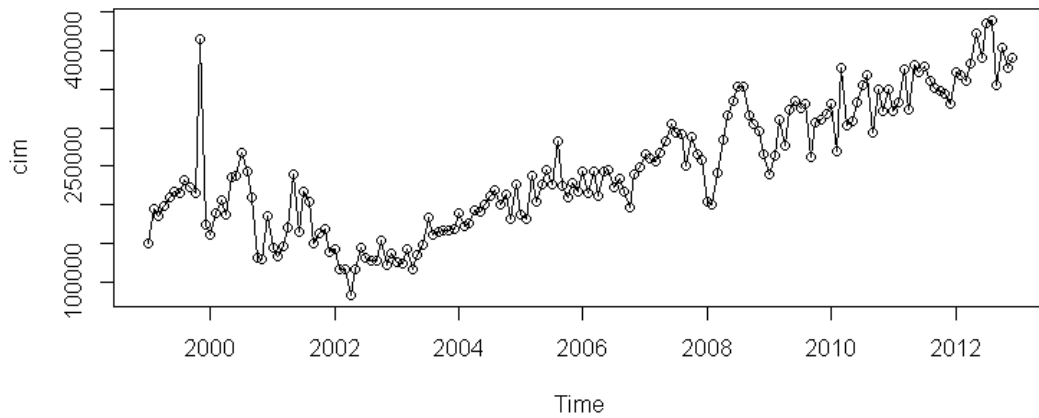
## **Appendix B : Time Series Plots**

**Figure(1): Quarterly Real GDP Of Palestine(\$)**



**Figure(2): Quarterly Employment Rate In Israel And Settlements(%)**



**Figure(3): Monthly CPI In Palestine****Figure(4): Monthly Cost Of Imports In Palestine(\$)**

### **Appendix C : Testing The Flat Aggregation Scheme(AGK test)**

When researchers follow a MIDAS regression approach, It is interested to examine whether estimating MIDAS regression does not differ from time-averaging aggregation. For example in Exponential Almon Lag weights, when all parameters ( $\theta$ 's) equal zero, this implies to use time-averaging since all weights will be equal. Andreou, Ghysels and Kourtellos(2007) suggested to use simple Lagrange multiplier chi square test to examine the  $H_0 : \theta = 0$ . The test statistic LM could be written as the following:

$$LM = n \frac{SSE_{TA} - SSE_M}{SSE_{TA}} \xrightarrow{d} \chi^2(r)$$

Where  $SSE_{TA}$  is the sum of squared errors from the linear regression based on Time-Averaging method, and  $SSE_M$  is the sum of squared errors based on MIDAS regression,  $r$  is the number of aggregated high-frequency variables and  $n$  is the number of low-frequency observations, this test also called Andreou, Ghysels, Kourtellos(AGK) LM test[32].

## Appendix D : The R-Codes Used In Analysis

```

*****plotting figure(1) in section 3.1*****
data("USunempr")
data("USrealgdp")
USunempr
USrealgdp
plot(USrealgdp, main="Figure(1): Yearly GDP and Monthly Unemployment Rates")
plot(USunempr)
*****plotting figure(3) in section 3.5*****
library(ggplot2)
library(reshape2)
library(midasr)
y1=nealmon(c(1,0.01,-0.0025), 20, m)
y2=nealmon(c(1,0.01,-0.0099), 20, m)
y3=nealmon(c(1,0.099,-0.0099), 20, m)
t=1:20
df <- data.frame(t,y1,y2,y3)
ggplot(df, aes(t)) + # basic graphical object
  geom_line(aes(y=y1), colour="red", size=(1)) + # first layer
  geom_line(aes(y=y2), colour="green", size=(1))+
  geom_line(aes(y=y3), colour="black", size=(1))+
  ylab("Weight on Lag")+ xlab("Lag")+
  ggtitle("Figure(3): Exponential Almon Polynomial Weighting
Function")
#plot(y1) #may need to execute before the next lines
legend("topright", title = "parameter values",
legend=c("theta1=0.01, theta2=-0.0025",
"theta1=0.01, theta2=-0.0099",
"theta1=0.099, theta2=-0.0099"), # puts text in the legend
lwd=c(2,2,2),col=c("red","green","black"))
*****Calling Libraries*****
library(TSA)
library(tseries)
library(zoo)
library(lmtest)
library(MASS)
library(numDeriv)
library(sandwich)
library(strucchange)
library(urca)
library(vars)
library(midasr)

```

```

#####Calling Data and Checks#####
th1 = read.csv("qd.csv", header=T, sep=";" )
attach(th1)
th2 = read.csv("md.csv", header=T, sep=";" )
attach(th2)
gdp=th1$gdp
eii=th1$eii
cpi=th2$cpi
cim=th2$cim
b1=BoxCox.ar(gdp)
b1$ci# 1
b1=BoxCox.ar(eii)
b1$ci # -0.5 close to 0
b1=BoxCox.ar(cpi)
b1$ci# 0 i.e. log
b1=BoxCox.ar(cim)
b1$ci# 0.3 close to zero
lgdp=log(gdp)
leii=log(eii)
lcpi=log(cpi)
lcim=log(cim)
adf.test(lgdp, k=4)# not stationary
adf.test(leii)# not stationary
adf.test(lcpi)# not stationary
adf.test(lcim)# not stationary
adf.test(diff(lgdp))# stationary
adf.test(diff(leii))# stationary
adf.test(diff(lcpi))# stationary
adf.test(diff(lcim))# stationary
lgdpfl=zlag(lgdp)
set.seed(1001)
## Number of low-frequency observations
n <- 56
## Linear trend and higher-frequency explanatory variables
## (quarterly and monthly)
trend <- c(1:n)
## *****Fitting RMIDAS*****
eq.r <- midas_r(lgdp ~ trend + lgdpfl + leii +
mles(lcpi, 0:2, 3, nealmon) +
mles(lcim, 0:2, 3, nealmon),
start = list(lcpi = c(0, 0, 0),lcim = c(0, 0, 0)))
summary(eq.r)
r=resid(eq.r)

```

```

adf.test(r, k=2)# stationary
pp.test(r)
midas_coef(eq.r)
agk.test(eq.r)# the test accept H0 i.e. all thetas are zero
hAh.test(eq.r)#the test rject H0 i.e. the restriction not hold
# i.e. wi(weights) not equal b(k;theta)(nealmon)
#starting values taken from the preveous summary
eq.r <- midas_r(lgdp ~ trend + lgdpfl + leii +
mls(lcpi, 0:2, 3, nealmon) +
mls(lcim, 0:2, 3, nealmon),
start = list(lcpi = c(-5.995e-01, -3.161e-01, -1.544e+00),
lcim = c(3.095e-01, 2.217e+00, -5.910e-01)))
summary(eq.r)
r=resid(eq.r)
adf.test(r, k=2)# stationary
pp.test(r)
midas_coef(eq.r)
agk.test(eq.r)# the test accept H0 i.e. all thetas are zero
hAh.test(eq.r)#the test rject H0 i.e. the restriction not hold
# i.e. wi(weights) not equal b(k;theta)(nealmon)
eq.r$opt
## *****Fitting UMIDAS*****
eq.u <- lm(lgdp ~ trend + lgdpfl + leii +
mls(lcpi, k = 0:2, m = 3) +
mls(lcim, k = 0:2, m = 3))
summary(eq.u)
r=resid(eq.u)
adf.test(r, k=2)
pp.test(r)
# the same as lm
eq.u3 <- midas_u(lgdp ~ trend + lgdpfl + leii +
mls(lcpi, k = 0:2, m = 3) +
mls(lcim, k = 0:2, m = 3))
summary(eq.u3)
#reduced form cpi eleminated
eq.u <- lm(lgdp ~ trend + lgdpfl + leii +
mls(lcim, k = 1:1, m = 3))
summary(eq.u)
r=resid(eq.u)
adf.test(r, k=2)
pp.test(r)

## *****Fitting Time-Averaging Model*****

```

```

th= read.csv("ffd.csv", header=T, sep=";")
attach(th)
gdp=th$gdp
eii=th$eii
cpi=th$cpi
cim=th$cim
lgdp=log(gdp)
leii=log(eii)
lcpi=log(cpi)
lcim=log(cim)
# all these variables were tested and are I(1)
set.seed(1001)
## Number of low-frequency observations
n <- 56
## Linear trend and higher-frequency explanatory variables
## (quarterly and monthly)
trend <- c(1:n)
lgdpfl=zlag(lgdp)
eq.u <- lm(lgdp ~ trend + lgdpfl + leii + lcpi + lcim)
summary(eq.u)
# reduced form cpi eliminated
eq.u <- lm(lgdp ~ trend + lgdpfl + leii + lcim)
summary(eq.u)
r=resid(eq.u)
adf.test(r, k=2)
pp.test(r)
## *****Cointegration Tests*****
cim2=th$cim2
adf.test(gdp)
adf.test(eii)
adf.test(cpi0)
adf.test(cpi1)
adf.test(cpi2)
adf.test(cim0)
adf.test(cim1)
adf.test(cim2)
adf.test(diff(gdp))
adf.test(diff(eii))
adf.test(diff(cpi0))
adf.test(diff(cpi1))
adf.test(diff(cpi2))
adf.test(diff(cim0))
adf.test(diff(cim1))

```

```

adf.test(diff(cim2))
data2=cbind(gdp,eii,cpi0,cpi1,cpi2,cim0,cim1,cim2)
summary(ca.jo(data2,type="trace",ecdet="const"))
summary(ca.jo(data2,type="eigen",ecdet="const"))
#results at most 3 cointegration equations
# results after eliminating cpi0, cpi1, cpi2
data2=cbind(gdp,eii,cim0,cim1,cim2)
summary(ca.jo(data2,type="trace",ecdet="const"))
summary(ca.jo(data2,type="eigen",ecdet="const"))
## *****Fitting ECM Models*****
th3 = read.csv("diffqd.csv", header=T, sep=";" )
attach(th3)
th4 = read.csv("diffmd.csv", header=T, sep=";" )
attach(th4)
dlgdp=th3$dlgdp
flut=th3$flut
lgdpfl=th3$lgdpfl
dleii=th3$dleii
dlcpi=th4$dlcpi
dlcim=th4$dlcim
#rmidas ECM
eq.r <- midas_r(dlgdp ~ flut + lgdpfl + dleii +
m1s(dlcpi, 0:2, 3, nealmon) +
m1s(dlcim, 0:2, 3, nealmon),
start = list(dlcpi = c(0, 0, 0),
dlcim = c(0, 0, 0)))
summary(eq.r)
midas_coef(eq.r)
agk.test(eq.r)# the test accept H0 i.e. all thetas are zero
hAh.test(eq.r)#the test rject H0 i.e. the restriction not hold
# i.e. wi(weights) not equal b(k;theta)(nealmon)
eq.r$opt
# full ECM by U-MIDAS
eq.u <- midas_u(dlgdp ~ flut + lgdpfl + dleii +
m1s(dlcpi, 0:2, 3, nealmon) +
m1s(dlcim, 0:2, 3, nealmon))
summary(eq.u)
# reduced ECM by U-MIDAS
eq.u <- midas_u(dlgdp ~ flut +
m1s(dlcim, 1:2, 3, nealmon))
summary(eq.u)

# full ECM by time-averaging

```



```

th5 = read.csv("ffd_ecm.csv", header=T, sep=";" )
attach(th5)
dlgdp=th5$dlgdp
flut=th5$flut
lgdpfl=th5$lgdpfl
dleii=th5$dleii
dlcpi=th5$dlcpi
dlcim=th5$dlcim
eq.a <- lm(dlgdp ~ flut + lgdpfl + dleii +
dlcpi + dlcim)
summary(eq.a)
# reduced ECM by time-averaging
eq.a <- lm(dlgdp ~ flut +
dlcim)
summary(eq.a)
## *****Forecasting*****
th1 = read.csv("qd.csv", header=T, sep=";" )
attach(th1)
th2 = read.csv("md.csv", header=T, sep=";" )
attach(th2)
gdp=th1$gdp
eii=th1$eii
cpi=th2$cpi
cim=th2$cim
lgdp=log(gdp)
leii=log(eii)
lcpi=log(cpi)
lcim=log(cim)
lgdpfl=zlag(lgdp)
fulldata <- list(lgdp=lgdp, leii=leii, lcpi=lcpi, lcim=lcim,
cim=cim, lgdpfl=lgdpfl, trend=1:56)
avgf <- average_forecast(list(eq.u, eq.r, eq.ta),
data=fulldata,
insample=1:52,outsample=53:56,
type="fixed",
measures=c("MSE","MAPE","MASE"),
fweights=c("EW","BICW","MSFE","DMSFE"))
avgf
newdata <- list(leii=leii[56], lcim=c(lcim[54],lcim[55],lcim[56]),
lgdpfl=lgdp[56], trend=57)
nf=forecast(eq.u, newdata = newdata)
exp(nf)

```

## **References**

- [01]. Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov. *The MIDAS Touch: Mixed Data Sampling Regression Models*. University of North Carolina & The Anderson School UCLA, June 22, 2004.
- [02]. Sinko, Arthur. *Some Applications of Mixed Data Sampling Regression Models*. Chapel Hill: Arthur Sinko, 2008.
- [03]. Foroni, Claudia. *Econometric Models for Mixed frequency-data*. Florence: European University Institute/Department of Economics, 2012.
- [04]. Wohlrabe, Klaus. *Forecasting with Mixed-frequency Time Series Models*. University of Munich, 2008.
- [05]. Clements, Michael P., and Ana Beatriz Galvao. *Microeconomic Forecasting with Mixed Frequency Data: Forecasting US Output Growth*. No. 616. working paper. University of London, October 2007.
- [06]. Foroni, Claudia, Massimiliano Marcellino, and Christian Schumacher. *U-MIDAS: MIDAS regressions with unrestricted lag polynomials*. Deutsche Bundesbank, 2011.
- [07]. Foroni, Claudia, and Massimiliano Marcellino. *A Survey of Econometric Methods for Mixed-Frequency Data*. Norges Bank, February 6, 2013.
- [08]. Andreou, Elena, Eric Ghysels, and Andros Kourtellis. "Forecasting with mixed-frequency data." In *Oxford Handbook on Economic Forecasting*. European Research Council, 2010.
- [09]. Chen, Yu-chin, and Wen Jen Tsay. *Forecasting Commodity Prices with Mixed-Frequency Data: An OLS-Based Generalized ADL Approach*. University of Washington & Academia Sinica, April 2011.
- [10]. Ghysels, Eric. *MIDAS Regressions and Their Applications in Finance and Macroeconomics*. SAMSI, September 2005.
- [11]. Kutner, Michael H., Christopher J. Nachtsheim, John Neter, and William Li. *Applied Linear Statistical Models*. Fifth Edition. McGraw-Hill/Irwin, 2005.
- [12]. Pankratz, Alan. *Forecasting With Univariate Box- Jenkins Models: CONCEPTS AND CASES*. John Wiley & Sons. Inc, 1983.

- [13]. Chatfield, Chris. *The Analysis Of Time Series: An Introduction*. Fifth Edition. University of Bath: Chapman & Hall/CRC, 1995.
- [14]. Harvey, Andrew C. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge University Press, 1990.
- [15]. Sinko, Arthur, Michael Sockin, and Eric Ghysels. *Matlab Toolbox for Mixed Sampling Frequency Data Analysis using MIDAS Regression Models*. 1.0. Federal Reserve Bank of New York, January 5, 2011.
- [16]. Ghysels, Eric. *Matlab Toolbox for Mixed Sampling Frequency Data Analysis using MIDAS Regression Models*. 5.0. Federal Reserve Bank of New York, May 29, 2013.
- [17]. Shi, Wensi. *MIDAS Predicting Volatility at Different Frequencies*. Uppsala University, May 18, 2010.
- [18]. Armesto, Michelle T., Kristie M. Engemann, and Michael T. Owyang. "Forecasting with Mixed Frequencies." *St. Louis*, November/December 2010: 36-521.
- [19]. Asimakopoulos, Stylianos, Joan Paredes, and Thomas Warmedinger. *Forecasting fiscal time series using mixed frequency data*. European Central Bank, May 2013.
- [20]. De Jesus, Miguel Karlo Reyes. *Data Pre-cleaning Through Exponential Smoothing as Antecedent to a Robustified Mixed-Data Sampling (MIDAS) Regression*. University of Bielefeld, 2010.
- [21]. Goetz, Thomas B., Alain Hecq, and Jean Pierre Urbain. *Forecasting Mixed Frequency Time Series with ECM-MIDAS Models*. Maastricht Universit, SBE, Department of Quantitative Economics, December 16, 2011.
- [22]. Clements, Michael P., and Ana Beatriz Galvão. *Macroeconomic Forecasting with Mixed Frequency Data: Forecasting US output growth and inflation*. WARWICK ECONOMIC RESEARCH PAPERS. University of Warwick, July 2006.
- [23]. Khim, Venus, and Sen Liew. "Which Lag Length Selection Criteria Should We Employ?" September 17, 2004: 1-9.
- [24]. Hurvich, Clifford M. "Bias of the corrected AIC criterion for underfitted regression and time series models." *Biometrika*, 1990: 499-509.
- [25]. Dahyot, Rozenn. *Time series and Applied Forecasting*. Trinity College Dublin, 2013.

- [26]. Ghysels, Eric, Jonathan B. Hill, and Kaiji Motegi. *Testing for Granger Causality with Mixed Frequency Data*. Federal Reserve Bank of New York/European Central Bank, September 6, 2013.
- [27]. Ghysels, Eric, and Isaac J. Miller. *Testing for Cointegration with Temporally Aggregated and Mixed-frequency Time Series*. University of North Carolina at Chapel Hill/University of Missouri, July 1, 2013.
- [28]. Miller, Isaac J. *Mixed-frequency Cointegrating Regressions with Parsimonious Distributed Lag Structures*. Department of Economics, University of Missouri, 2013.
- [29]. Miller, Isaac J. *Cointegrating MiDaS Regressions and a MiDaS Test*. Department of Economics, University of Missouri, June 14, 2011.
- [30]. Seonga, Byeongchan, Sung K. Ahnb, and Peter A. Zdrozny. *Estimation of Vector Error Correction Models with Mixed-Frequency Data*. Chung-Ang University/Washington State University/Bureau of Labor Statistics, 2 Massachusetts Ave., October 2, 2008.
- [31]. Ghysels, Eric, Arthur Sinkoz, and Rossen Valkanov. *MIDAS Regressions: Further Results and New Directions*. University of North Carolina/UCSD, February 7, 2006.
- [32]. Andreou, Elena, Eric Ghysels, and Andros Kourtellos. *Regression Models with Mixed Sampling Frequencies*. Nicosia & Chapel Hill, March 16, 2009.
- [33]. Seong, Byeongchan, Sung K. Ahn, and Peter A. Zdrozny. *Cointegration Analysis With Mixed-Frequency Data*. CESifo Working Papers, March 2007.
- [34]. Ghysels, Eric. *Matlab Toolbox for Mixed Sampling Frequency Data Analysis using MIDAS Regression Models*. 8.0. Federal Reserve Bank of New York, October 9, 2013.
- [35]. Aref, Mohammed, and Michel Dombrecht. *A Short Run Reduced Form Equation For Real Gdp Growth In Opt*. Ramallah: PMA(Palestine Monetary Authority), April 2012.
- [36]. Khalil, Saed, and Michel Dombrecht. *The Autoregressive Itributed Lag Approach To Co-Integration Testing: Application To Opt Inflation*. Ramallah: Palestine Monetary Authority(PMA), November 2011.
- [37]. Khalil, Saed, and Michel Dombrecht. *Estimation Of The Equilibrium Real Exchange Rate In The Occupied Palestinian Territory*. Ramallah: PMA, October 2011.
- [38]. Authority, Palestine Monetary. *Inflation Report*. Inflation Report, PMA, Palestine Monetary Authority, 2013.

- [39]. Guay, Alain, and Alain Maurin. "An adaptation of the MIDAS regression model for estimating and forecasting quarterly GDP: Application to the case of Guadeloupe." April 2008.
- [40]. Marsilli, Clément. "Variable Selection in Predictive MIDAS Models." November 2013.
- [41]. Joyeux, Roselyne, and George Milunovich. Encyclopedia of Life Support Systems (EOLSS), 2011.
- [42]. Shumway, Robert H., and David S. Stoffer. *Time Series Analysis And Its Applications With R Examples*. Third Edition. Springer Texts in Statistics, 2011.
- [43]. Cryer, Jonathan D., and Kung Sik Chan. *Time Series Analysis With Applications in R*. Second Edition. Springer Texts in Statistics, 2008.
- [44]. Nielsen, Heino Bohn. "Generalized Method of Moments (GMM) Estimation." *Econometrics 2*. 2005.
- [45]. Hansen, Lars Peter. *Generalized Method of Moments Estimation*. National Science Foundation & University of Chicago, June 17, 2007.
- [46]. Arellano, Manuel. *Generalized Method of Moments and Optimal Instruments*. Manuel Arellano, February 26, 2009.
- [47]. Gavin, Henri P. "The Levenberg-Marquardt method for nonlinear least squares: curve-fitting problems." *Duke University*. October 9, 2013. [www.duke.edu/~hpgavin](http://www.duke.edu/~hpgavin).
- [48]. Gratton, S., A.S. Lawless, and N.K. Nichols. *Approximate Gauss-Newton methods for nonlinear least squares problems*. Numerical Analysis Report, UK: The University of Reading.
- [49]. Kuan, Chung Ming. *Introduction To Econometric Theory*. Institute of Economics, Academia Sinica, 2000.
- [50]. Hansen, Bruce E. *ECONOMETRICS*. University of Wisconsin, Department of Economics, 2014.
- [51]. Parker, Jeffrey. *Theory and Practice of Econometrics*. REED College. 2010. <http://academic.reed.edu/economics/parker/s10/312/readings.html#s13>.
- [52]. Kirchgässner, Gebhard, and Jürgen Wolters. *Introduction To Modern Time Series Analysis*. Springer, 2007.

[53]. Golinelli, Roberto. *Lectures on modelling non stationary time series*. Department of Economics, Strada Maggiore, June 2005.

[54]. Ssekuma, Rajab. *A STUDY OF COINTEGRATION MODELS WITH APPLICATIONS*. UNIVERSITY OF SOUTH AFRICA, June 2011.

## Index

### A

ADF, 19, 20, 21, 22, 24, 49, 53  
 ADL, 10, 11, 12, 25, 26, 30, 38  
 aggregating, 3  
 Aggregation, 4, 9, 10, 16, 58  
 AGK, 43, 51, 53, 58  
 AGK test, 43, 44  
 AIC, 18  
 Almon, 3, 14, 15, 16, 17, 25,  
 26, 43, 48, 58, 59  
 ARIMA, 1, 3  
 ARMA, 29  
 autoregressive distributed lag,  
 4, 10, 11, 12, 25

### B

Beta, 3, 16  
 Beta function, 16  
 BIC, 18

### C

causality, 30, 48  
 CIM, 1, 2, 3, 9, 40, 41, 43, 48  
 coefficient, 10, 11, 20, 22, 23,  
 27, 39, 42, 45, 50, 51, 52  
 coefficients, 10, 11, 13, 14, 27,  
 28, 29, 30, 42, 47, 48  
 cointegrating relationship, 22,  
 25  
 Cointegration, 19, 21, 25, 49,  
 62  
 Consumer Price Index, 6, 40,  
 48  
 cost of imports, VIII, 1, 5, 7, 8,  
 40, 41, 43, 44, 45, 46, 48  
 CPI, 40, 41, 43, 48  
 Criteria, 18  
 cumulative, 27, 28

### D

dynamic, 5, 12, 26, 27, 28, 29,  
 34

### E

ECM, 22, 25, 26, 41, 42, 43,  
 44, 45, 53, 54, 55, 63, 64  
 economic, VIII, 1, 2, 25, 28, 34,  
 35, 40, 47  
 EII, 41, 43, 45, 49  
 employment rate in Israel and  
 settlements, 41, 45, 46  
 equilibrium, 5, 21, 26, 41, 42,  
 44, 45, 48  
 error, 3, 4, 5, 18, 19, 20, 21, 22,  
 25, 29, 30, 34, 35, 36, 39,  
 47, 50, 51, 52, 53, 54, 55  
 error correction, 22, 45  
 Exponential, 3, 14, 15, 16, 25,  
 26, 43, 48, 58, 59

### F

forecasting, VIII, 3, 4, 5, 8, 9,  
 11, 14, 25, 38, 40, 47, 48  
 forecasts, 5, 8, 38, 40, 46

### G

GDP, VIII, 1, 2, 3, 5, 6, 7, 8, 9,  
 40, 41, 43, 44, 45, 46, 47,  
 48, 49, 59  
 GMM, 17, 34  
 Gross Domestic Product, 1, II,  
 5, 6, 47

### H

higher-frequency, 8, 9, 10, 11,  
 12, 13, 15, 18, 60, 62

### I

independent, 1, 2, 6, 10, 18, 19,  
 23, 30, 38, 39, 41

### L

lag, VIII, 3, 4, 5, 6, 7, 10, 11,  
 13, 16, 18, 20, 25, 26, 27,  
 28, 29, 30, 41, 42, 43, 44,  
 45, 46, 48  
 leads, 5, 8, 11, 14, 23, 30, 37  
 likelihood, 17, 22, 24, 32, 33,  
 34  
 long-run, VIII, 5, 21, 22, 25, 26,  
 28, 29, 41, 42, 43, 45, 46,  
 47, 48  
 lower-frequency, 7, 9, 10, 13

### M

MAPE, 39, 46, 47, 64  
 MIDAS, VIII, 3, 4, 5, 6, 7, 8, 9,  
 11, 12, 13, 14, 16, 17, 18,  
 25, 26, 30, 32, 38, 39, 40,  
 41, 42, 43, 44, 45, 46, 47,  
 48, 50, 51, 52, 53, 58, 63  
 multicollinearity, 30, 31

### N

NLS, 4, 16, 17, 36, 48, 51, 53  
 nonstationarity, 19

### O

OLS, 4, 17, 21, 22, 27, 30, 35,  
 36, 38, 41, 48

### P

Palestine Monetary Authority,  
 40

parameter, 11, 12, 32, 34, 37,  
50, 51, 52, 53, 54, 55, 59  
period, 7, 8, 11, 12, 14, 19, 25,  
26, 27, 28, 41, 48  
PMA, 40, 41, 48

## R

Regression, 1, II, 1, 12, 50, 51,  
52  
R-MIDAS, VIII, 6, 12, 13, 25,  
43, 44, 47

## S

short-run, VIII, 5, 22, 25, 26,  
41, 42, 43, 44, 45, 48

stationarity, 19, 21, 23, 44  
step-weighting, 6, 11

## T

time series, VIII, 1, 2, 3, 10, 18,  
19, 23, 25, 33, 38, 43, 48  
time-averaging, 6, 10, 43, 58,  
63, 64

## U

U-MIDAS, VIII, 4, 5, 11, 25,  
42, 43, 44, 46, 47  
unit root, 19, 20, 21  
Unit Root, 19

## V

VAR, 3, 23, 24  
variance, 3, 6, 10, 18, 19, 30,  
31, 32, 38, 39, 41, 47

## W

white noise, 10, 11, 12, 13, 19,  
20, 25, 31, 32, 33, 35, 38,  
39, 41