
Short Dynamic Panel Data Models in the Presence of a Higher Frequency Regressor(s)

Alaaeldin M. Dyab¹, Sayed M. ElSayed², Amal A. Fattah² *

¹ Department of Quantitative Methods, Faculty of Commerce, South Valley University, 83523, Egypt

² Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, 12613, Egypt

* **Correspondence:** asoubh84@cu.edu.eg

Abstract: The study motivates the estimation problem of short dynamic panel data models in the presence of higher frequency regressors where it is required to weigh the regressor series to be at the same frequency of the dependent variable. With the focus on First Difference Generalized Method of Moments (DIF GMM) and System Generalized Method of Moments (SYS GMM) estimators, short dynamic panel data models were discussed extensively in two cases the first is the balanced case and the second is the mixed frequency case where there is a higher frequency regressor(s). The weighting of the higher frequency regressor can be carried out using Mixed Data Sampling (MIDAS). In the presence of a higher frequency regressor and by assuming AR (1) of the lower frequency time dimension as a data generating process for the regressor, the initial conditions satisfying mean stationarity required for SYS GMM MIDAS were derived. Monte Carlo simulations were carried out where the results showed that using SYS GMM MIDAS exploited more power than DIF GMM MIDAS. In addition, using the average for weighting as well as testing the significance of the higher frequency regressor parameter resulted in low power.

Keywords: Short Dynamic Panel, Mixed Frequency, Mixed Data Sampling (MIDAS), Generalized Method of Moments.

Mathematics Subject Classification: 94A20; 20K21; 90C56; 78M05.

Received: 30 August 2023; Revised: 20 September 2023; Accepted: 15 October 2023; Online: 20 October 2023.



Copyright: © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The panel data set has two dimensions one is cross-sectional and the other is for time that makes it more informative and gives accurate inference of model parameters. In addition, it can be used

to model dynamic processes by adding the lagged values of the dependent variable, the explanatory variable(s) or both. For short dynamic panel data models where the number of cross-sectional units (N) is higher than the number of time points (T) i.e., $N > T$, dynamics can be accounted for by adding a lagged dependent variable however the dependent variable becomes a function of the individual specific effects. This endogeneity problem makes the least-squares estimator inconsistent. Using estimators based on generalized method of moments estimator can produce more consistent estimates such as the First Difference GMM (FID GMM) introduced by [1] and System GMM (SYS GMM) proposed by [2].

One kind of unbalanced data is when the frequency of the regressor(s) is higher than the frequency of the dependent variable known as mixed frequency e.g., weekly vs. daily, quarterly vs. monthly, or yearly vs. monthly data, . . . etc. Mixed frequency data is one of the most challenging problems in econometrics where it is required to aggregate the higher frequency regressor to have the same frequency as the dependent variable. The problem with temporal aggregation is the anticipated loss of information such that the dynamics of the aggregate model becomes different from that in the unbalanced data set i.e., specification bias. One way to deal with mixed frequency data is using mixed data sampling techniques (MIDAS) introduced by [3] where a parsimonious lag polynomials are used to weight the higher frequency regressor(s) series preventing parameters proliferation. The advantage of MIDAS is that it preserves the data generating process as possible with fewer number of parameters required to estimate the model. In case of small difference between the frequency of the dependent variable and the regressor(s) e.g., quarterly vs. monthly data, the unrestricted MIDAS (U-MIDAS) proposed by [4] can be used where no conditions are imposed on the weight parameters. Applying MIDAS to short dynamic panel with mixed frequency data was introduced by [5] where a two parameters exponential Almon lag function is used to parametrize the higher frequency regressor series while the First Difference GMM serves as estimation method.

The estimation is carried out numerically where a grid search is constructed to make inference about the exponential Almon lag parameters by inverting the Sargan test of overidentifying restrictions such that the set of all estimates with Sargan test p-values greater than a given level of significance form a confidence interval for the model parameters from which the one with highest p-value is regarded as point estimate whereas an empty set signals a lack of fit. This study relies on the approach used in [5] to parameterize the higher frequency regressor series but it applies the SYS GMM as estimation method alongside with DIF GMM while comparing between these two methods using Monte Carlo simulation in the sense of their power and empirical size.

The study has the following structure: section 2 provides a literature review about short dynamic panel data models in balanced and mixed frequency cases. With focus on the first difference and system GMM estimators, balanced short dynamic panel data models were extensively discussed in section 3, while including a higher frequency regressor(s) in short dynamic panel data models was represented in section 4. A simulation study was conducted and illustrated in section 5. Finally, concluding remarks were presented in section 6.

2. Literature Review

Reference [6] put forth an Ordinary Least Squares (OLS) estimator to address short dynamic panel data issues. However, the OLS estimator becomes biased and inconsistent due to the endogeneity

problem. To tackle this, they transformed the model using a within transformation, applying the Least Squares Dummy Variables (LSDV) estimator, which still exhibits bias and consistency reliance on a large T value. They also introduced a Maximum Likelihood (ML) estimator, but its consistency depends on initial conditions. They later introduced the Instrumental Variables (IV) estimator after differencing the model, which is consistent regardless of initial conditions but not necessarily efficient. [7] introduced the Limited Information Maximum Likelihood (LIML) to address inconsistency of variance that arises in applying maximum likelihood to short dynamic panel data. [1] suggested enhancing the IV approach with additional instruments using the generalized method of moments, known as First Difference GMM (DIF GMM), and proposed three specification tests to ensure exogeneity. [8] introduced a method involving IV or GMM techniques to estimate the inverse covariance matrix, leading to a transformed model. [9] suggested a non-singular transformation and GMM method, which aligns with LIML. [10] found limitations in DIF GMM and proposed adding nonlinear moment conditions for efficiency, while [11] presented a bias-corrected DIF GMM estimator.

Reference [2] addressed weak instruments by using lagged differences alongside levels as instruments. [12] introduced Symmetrically Normalized GMM and LIML estimators, along with an overidentification restrictions test. [13] introduced transformed maximum likelihood for consistency. [3] introduced the MIDAS technique for mixed frequency data, using various lag polynomials for higher frequency regressors. [14] proposed GLS and GMM estimators for common time effects in dynamic panel data. [15] applied Quasi Maximum Likelihood (QML) method with time series heteroskedasticity, and [16] used a factor analytical approach to handle incidental parameters. [17] proposed an ssLIML estimator with better finite sample performance than GMM. [19] introduced TML and RML estimators, warning of potential issues with inference based on maximum likelihood for small N or T. [5] suggested MIDAS with mixed frequency data and different GMM estimators, including a parameterized exponential almon lag function. [20] applied the MIDAS approach to daily satellite data using a Beta lag parameterization, outperforming daily averages in predicting monthly industrial production.

For robust estimation methods: Alvarez et al. [21] identified biases in within group OLS, one-step GMM, and LIML estimators, and showed that GMM and LIML become asymptotically equivalent with increasing T. They also proposed a heteroskedasticity-robust MLE estimator. Almongy et al. [22] discussed robust estimation methods to estimate parameters of distributions models. Reference [23] discussed comparison between M estimation, S estimation, and MM estimation methods of robust estimation with application and simulation. Alshenawy et al. [24] discussed OLS and GLM based on robust estimation method to study of statistical and intelligent classification models for prediction diabetes. Kamel et al. [25] discussed some computational methods by R program.

3. Balanced Short Dynamic Panel Data Model

This section is intended to illustrate the estimation of balanced short dynamic panel data models using the first difference GMM (DIF GMM) and system GMM (SYS GMM). When the number of cross-sectional units (N) is sufficiently large while the time dimension (T) is limited, the model expressing these data can be written in the form of panel AR (1) as follows:

$$y_{it} = \lambda y_{i,t-1} + x_{it} \beta + \mu_{it} \quad (3.1)$$

$$u_{it} = \mu_i + v_{it} \quad (3.2)$$

Where $t = 2, \dots, T$, $i = 1, \dots, N$, λ is a scalar, x_{it} is $1 \times K$ vector of K observations for the explanatory variables concerning specific individual (i) and time (t), β is a $K \times 1$ vector of unknown coefficients, and u_{it} is the error term, μ_i unobservable individual specific effect, and v_{it} is idiosyncratic error that differs within individuals and throughout time.

Under the following assumptions:

$$A1. |\lambda| < 1.$$

$$A2. \mu_i \sim \text{IID}(0, \sigma_\mu^2).$$

$$A3. v_{it} \sim \text{IID}(0, \sigma_v^2).$$

$$A4. E[\mu_i x_{it}] \neq 0.$$

$$A5. E[v_{it} \mu_i] = 0$$

$$A6. E[v_{it} y_{it-j}] = 0; j = 1, \dots, t-1.$$

A7. x_{it} is partitioned to $(x_{it}^{(1)}, x_{it}^{(2)})$ where $x_{it}^{(1)}$ is $1 \times K_1$ vector of strictly exogenous explanatory variables such that $E[x_{it}^{(1)'} v_{is}] = 0_{k_1} \forall t, s; s = 1, \dots, T$ and $x_{it}^{(2)}$ is $1 \times K_2$ vector of predetermined explanatory variables satisfying $E[x_{it}^{(2)'} v_{is}] = 0_{k_2}, t \leq s$.

The model in (3.1) can be rewritten in matrix form as:

$$y = y_{-1} + X\beta + u \quad (3.3)$$

Where $y = [y_{12}, \dots, y_{N2}, \dots, y_{1T}, \dots, y_{NT}]$ is $N(T-1) \times 1$ vector of observations for the dependent variable, $y_{-1} = [y_{11}, \dots, y_{N1}, \dots, y_{1,T-1}, \dots, y_{N,T-1}]$ is $N(T-1) \times 1$ vector of lagged observations for the dependent variable, X is $N(T-1) \times K$ matrix of observations for the explanatory variables, and $[u_{12}, \dots, u_{N2}, \dots, u_{1T}, \dots, u_{NT}]$ is $N(T-1) \times 1$ vector of error terms that can be expressed using matrix form as:

$$u = z_\mu \mu + v$$

Where $z_\mu = I_N \otimes l_{T-1}$ is $N(T-1) \times N$ matrix, I_N is identity matrix of order N , l_{T-1} is $(T-1) \times 1$ vector of ones, and $v = [v_{12}, \dots, v_{N2}, \dots, v_{1T}, \dots, v_{NT}]$ is a $N(T-1) \times 1$ vector of idiosyncratic error terms.

Furthermore (3.3) can be rewritten as:

$$y = Z\delta + u \quad (3.4)$$

Where $\delta' = [\lambda \ \beta']$ is $1 \times (K+1)$ vector of unknown coefficients, and $Z = [y_{-1} \ X]$ is $N(T-2) \times (K+1)$ matrix of observations for $(K+1)$ regressors.

The individual specific effects μ_i in (3.1) can be cancelled using the first difference transformation yielding:

$$\Delta y_{it} = \lambda \Delta y_{it-1} + \Delta x_{it} \beta + \Delta v_{it} \quad (3.5)$$

Where $i = 1, \dots, N$, $t = 1, \dots, T$.

The model in (3.5) can be written in matrix form as:

$$\Delta y = \lambda \Delta y_{-1} + \Delta X \beta + \Delta v \quad (3.6)$$

Where $\Delta y = [\Delta y_{13}, \dots, \Delta y_{1T}, \dots, \Delta y_{N3}, \dots, \Delta y_{NT}]$ is $N(T-2) \times 1$ vector of differenced observations for the dependent variable, $\Delta y_{-1} = [\Delta y_{12}, \dots, \Delta y_{1,T-1}, \dots, \Delta y_{N2}, \dots, \Delta y_{N,T-1}]$ is $N(T-2) \times 1$ vector

of lagged differenced observations for dependent variable, ΔX is $N(T-2) \times K$ matrix of differenced observations for K explanatory variables, $\Delta v = [\Delta v_{13}, \dots, \Delta v_{1T}, \dots, \Delta v_{N3}, \dots, \Delta v_{NT}]'$ is $N(T-2) \times 1$ vector of differenced idiosyncratic error terms.

Moreover (3.6) can be rewritten as:

$$\Delta y = \Delta Z \delta + \Delta v \quad (3.7)$$

Where $\Delta Z = [\Delta y_{-1} \ \Delta X]$ is $N(T-2) \times (K+1)$ matrix of differenced observations for $(K+1)$ regressors.

3.1. First Difference GMM

To get a consistent estimator for the model in (3.5), [1] tried to find instruments that are not correlated with the differenced error Δv_{it} ; assuming the assumptions from A1 to A7 are satisfied these instruments are:

When $t=3$:

$$y_{i3} - y_{i2} = \lambda (y_{i2} - y_{i1}) + (x_{i3}^{(1)} - x_{i2}^{(1)})\beta^{(1)} + (x_{i3}^{(2)} - x_{i2}^{(2)})\beta^{(2)} + (v_{i3} - v_{i2})$$

$y_{i1}, x_{i1}^{(1)}, x_{i2}^{(1)}, x_{i3}^{(1)}, x_{i1}^{(2)}, x_{i2}^{(2)}$ are valid instruments.

When $t = 4$:

$$y_{i4} - y_{i3} = \lambda (y_{i3} - y_{i2}) + (x_{i4}^{(1)} - x_{i3}^{(1)})\beta^{(1)} + (x_{i4}^{(2)} - x_{i3}^{(2)})\beta^{(2)} + (v_{i4} - v_{i3})$$

$y_{i1}, y_{i2}, x_{i1}^{(1)}, x_{i2}^{(1)}, x_{i3}^{(1)}, x_{i4}^{(1)}, x_{i1}^{(2)}, x_{i2}^{(2)}, x_{i3}^{(2)}$ are valid instruments.

As a general case, when $t = T$:

$$y_{it} - y_{i,t-1} = \lambda (y_{i,t-1} - y_{i,T-2}) + (x_{it}^{(1)} - x_{i,t-1}^{(1)})\beta^{(1)} + (x_{it}^{(2)} - x_{i,t-1}^{(2)})\beta^{(2)} + (v_{it} - v_{i,t-1})$$

$y_{i1}, \dots, y_{i,T-2}, x_{i1}^{(1)}, \dots, x_{it}^{(1)}, x_{i1}^{(2)}, \dots, x_{i,T-1}^{(2)}$ are valid instruments.

In case of predetermined explanatory variables:

When $t=3$:

$$y_{i3} - y_{i2} = \lambda (y_{i2} - y_{i1}) + (x_{i3}^{(2)} - x_{i2}^{(2)})\beta^{(2)} + (v_{i3} - v_{i2})$$

$y_{i1}, x_{i1}^{(2)}, x_{i2}^{(2)}$ are valid instruments.

When $t = 4$:

$$y_{i4} - y_{i3} = \lambda (y_{i3} - y_{i2}) + (x_{i4}^{(2)} - x_{i3}^{(2)})\beta^{(2)} + (v_{i4} - v_{i3})$$

$y_{i1}, y_{i2}, x_{i1}^{(2)}, x_{i2}^{(2)}, x_{i3}^{(2)}$ are valid instruments.

As a general case, when $t = T$:

$$y_{it} - y_{i,t-1} = \lambda (y_{i,t-1} - y_{i,T-2}) + (x_{it}^{(2)} - x_{i,t-1}^{(2)})\beta^{(2)} + (v_{it} - v_{i,t-1})$$

$y_{i1}, \dots, y_{i,T-2}, x_{i1}^{(2)}, \dots, x_{i,T-1}^{(2)}$ are valid instruments.

In DIF GMM, the matrix of valid instruments for each cross-sectional unit i can be partitioned into three sub-matrices as follows:

$$W_i = [W_i(y), W_i(x^{(1)}), W_i(x^{(2)})] \quad (3.8)$$

Such that:

$$W_i(y) = \begin{bmatrix} y_{i1} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & [y_{i1} & y_{i2}] & \dots & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & [y_{i1} & \dots & y_{i,T-2}] \end{bmatrix} \quad (3.9)$$

Where $W_i(y)$ is $(T-2) \times (T-2)(T-1)/2$ matrix.

$$W_i(x^{(1)}) = \begin{bmatrix} [x_{i1}^{(1)} & x_{i2}^{(1)} & x_{i3}^{(1)}] & 0_{k_1} & 0_{k_1} & 0_{k_1} & 0_{k_1} & \dots & 0_{k_1} & \dots & 0_{k_1} \\ 0_{k_1} & 0_{k_1} & 0_{k_1} & [x_{i1}^{(1)} & x_{i2}^{(1)} & x_{i3}^{(1)} & x_{i4}^{(1)}] & \dots & 0_{k_1} & \dots & 0_{k_1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0_{k_1} & 0_{k_1} & 0_{k_1} & 0_{k_1} & 0_{k_1} & 0_{k_1} & 0_{k_1} & \dots & [x_{i1}^{(1)} & \dots & x_{iT}^{(1)}] \end{bmatrix} \quad (3.10)$$

Where $W_i(x^{(1)})$ is $(T-2) \times (T-2)(T+3)K_1/2$ matrix.

$$W_i(x^{(2)}) = \begin{bmatrix} [x_{i1}^{(2)} & x_{i2}^{(2)}] & 0_{k_2} & 0_{k_2} & 0_{k_2} & \dots & 0_{k_2} & \dots & 0_{k_2} \\ 0_{k_2} & 0_{k_2} & [x_{i1}^{(2)} & x_{i2}^{(2)} & x_{i3}^{(2)}] & \dots & 0_{k_2} & \dots & 0_{k_2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \dots & \vdots \\ 0_{k_2} & 0_{k_2} & 0_{k_2} & 0_{k_2} & 0_{k_2} & \dots & [x_{i1}^{(2)} & \dots & x_{iT-1}^{(2)}] \end{bmatrix} \quad (3.11)$$

Where $W_i(x^{(2)})$ is $(T-2) \times (T-2)(T+1)K_2/2$ matrix.

For all cross-sectional units, the matrix of valid instruments in DIF GMM is:

$$W = [W'_1, \dots, W'_N]' \quad (3.12)$$

Where W is $N(T-2) \times P$ matrix.

The orthogonality conditions for the instrumental variables in the DIF GMM method are:

$$E[y_{i,t-s} \Delta v_{it}] = 0, \quad t = 3, \dots, T, \quad s = 2, \dots, T \quad (3.13)$$

$$E[x_{is}^{(1)'} \Delta v_{it}] = 0, \quad t = 3, \dots, T, \quad s = 1, \dots, T \quad (3.14)$$

$$E[x_{is}^{(2)'} \Delta v_{it}] = 0, \quad t = 3, \dots, T, \quad s = 1, \dots, T-1 \quad (3.15)$$

Alternatively, the orthogonality conditions in (3.13), (3.14) and (3.15) can be expressed using the matrix of valid instruments in (3.12) as:

$$M = W' \Delta v \quad (3.16)$$

Where M is $P \times 1$ vector.

Lemma 3.1: The first step DIF GMM estimator of the differenced model in (3.7) is:

$$\widehat{\delta}_1 = \left[\Delta Z' W (W (I_N \otimes G) W')^{-1} W' \Delta Z \right]^{-1} \left[\Delta Z' W (W (I_N \otimes G) W')^{-1} \right] \quad (\text{Arellano and Bond, 1991}).$$

Proof:

To apply GMM on (3.7), the following quadratic formula (Q) has to be minimized:

$$Q = M' A M \quad (3.17)$$

Where (A) is P×P positive definite weighting matrix.

Using (3.7) and substituting in (3.16) and (3.17), the quadratic formula in (3.17) can be rewritten as:

$$Q = [(\Delta y - \Delta Z \delta)' W] A [W' (\Delta y - \Delta Z \delta)] \quad (3.18)$$

Differentiating (3.18) with respect to δ and equating it with zero to obtain $\widehat{\delta}$ as:

$$\widehat{\delta} = (\Delta Z' W A W' \Delta Z)^{-1} \Delta Z' W A W' \Delta y \quad (3.19)$$

The weighting matrix in (3.17) might be taken as the inverse of the covariance matrix for moments conditions in (3.16) as follows:

$$A = [W' \text{cov}(\Delta v) W]^{-1} \quad (3.20)$$

Where $\text{cov}(\Delta v)$ is a $N(T-2) \times N(T-2)$ matrix that can be defined as:

$$\text{cov}(\Delta v) = \sigma_v^2 (I_N \otimes G) \quad (3.21)$$

Where $G = \begin{bmatrix} -2 & -1 & 0 & \dots & 0 & 0 & 0 \\ -1 & 2 & -1 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & -1 & 2 & -1 \\ 0 & 0 & 0 & \dots & 0 & -1 & 2 \end{bmatrix}$ is a $(T-2) \times (T-2)$ matrix.

Using (3.21), the weighting matrix (A) in (3.20) can be rewritten as:

$$A = \frac{1}{\sigma_v^2} [W' (I_N \otimes G) W]^{-1} \quad (3.22)$$

Using (3.22) and substituting in (3.19), the first step DIF GMM estimator can be defined as follows:

$$\widehat{\delta}_1 = [\Delta Z' W (W' (I_N \otimes G) W)^{-1} W' \Delta Z]^{-1} \Delta Z' W (W' (I_N \otimes G) W)^{-1} W' \Delta y \quad (3.23)$$

To get the second step DIF GMM estimator, the residuals of the first-step estimation $\widehat{\Delta v}$ is used to estimate the covariance matrix in (3.21) where $(W' (I_N \otimes G) W)$ in (3.23) is replaced with $V_N = \sum_{i=1}^N W_i' \Delta \widehat{v}_i \Delta \widehat{v}_i' W_i$ as follows:

$$\widehat{\delta}_2 = [\Delta Z' W V_N^{-1} W' \Delta Z]^{-1} \Delta Z' W V_N^{-1} W' \Delta y \quad (3.24)$$

where V_N is P×P matrix.

The covariance matrix of the second step DIF GMM estimator can be estimated as:

$$\widehat{\text{Cov}}(\widehat{\delta}_2) = [\Delta Z' W V_N^{-1} W' \Delta Z]^{-1} \quad (3.25)$$

3.2. System GMM

[2] noticed that DIF GMM might have poor finite sample bias especially in highly persistent series. It suggested using System GMM (SYS GMM) estimator relying on two simultaneous equations the first is the model based on levels in (3.1) and the second is the model based on first differences in (3.5).

Besides the orthogonality conditions defined in (3.13), (3.14), and (3.15) SYS GMM uses the following additional orthogonality conditions in case of strictly exogenous or predetermined explanatory variables:

$$E [\Delta y_{i,t-1} u_{it}] = 0, \quad t = 3, \dots, T \quad (3.26)$$

$$E [\Delta x_{i,t-1} u_{it}] = 0, \quad t = 3, \dots, T \quad (3.27)$$

In SYS GMM, the matrix of valid instruments for each cross-sectional unit i is defined as:

$$W_i^+ = \begin{bmatrix} W_i & 0 & 0_k & 0 & 0_k & \dots & 0 & 0_k \\ 0_p & [\Delta y_{i2} & \Delta x_{i2}] & 0 & 0_k & \dots & 0 & 0_k \\ 0_p & 0 & 0_k & [\Delta y_{i3} & \Delta x_{i3}] & \dots & 0 & 0_k \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_p & 0 & 0_k & 0 & 0_k & \dots & [\Delta y_{i,t-1} & \Delta x_{i,t-1}] \end{bmatrix}$$

Where 0_p and 0_k are $1 \times P$ and $1 \times K$ vectors of zeros respectively, W_i^+ is $2(T-2) \times P^+$ matrix, and $P^+ = (T-2)[(T+1) + (T+5)K]/2$.

The vector of differenced idiosyncratic error terms Δv defined in (3.6) and the last $(T-2)$ elements in the vector of error terms u defined in (3.3) can be stacked together as follows:

$$v_i^+ = [\Delta v_{i3}, \dots, \Delta v_{it}, u_{i3}, \dots, u_{it}]'$$

Where v_i^+ is a $2(T-2) \times 1$ vector.

For all cross-sectional units, the matrix of valid instruments in SYS GMM is:

$$W^+ = [W_1^+, \dots, W_N^+]' \quad (3.28)$$

Where W^+ is $2N(T-2) \times P^+$ matrix.

Similarly, the differenced idiosyncratic error terms and error terms v_i^+ can be defined for all cross-sectional units as:

$$v^+ = [v_1^+, \dots, v_N^+]'$$

Where v^+ is $2N(T-2) \times 1$ vector.

The orthogonality conditions for SYS GMM can be expressed using the matrix of valid instruments in (3.28) as follows:

$$M^+ = W^+ v^+ \quad (3.29)$$

Where M^+ is $P^+ \times 1$ vector.

The vector of differenced observations for the dependent variable Δy in (3.7) and the vector of observations for the dependent variable y in (3.4) can be stacked together as follows:

$$y^+ = [\Delta y' \quad y']'$$

Where y^+ is $2N(T-2) \times 1$ vector.

The matrix of differenced observations for the regressors ΔZ in (3.7) and the matrix of observations for the regressors Z in (3.4) can be stacked together as follows:

$$Z^+ = [\Delta Z' \quad Z']'$$

Where Z^+ is $2N(T-2) \times (K+1)$ matrix.

Lemma 3.2: The first step SYS GMM estimator of the model in (3.4) and the differenced model in (3.7) is:

$$\widehat{\delta}_1^+ = \left[Z^{+'} W^+ (W^+ (I_N \otimes G^+) W^{+'})^{-1} W^{+'} Z^+ \right]^{-1} Z^{+'} W^+ (W^+ (I_N \otimes G^+) W^{+'})^{-1} W^{+'} y^+$$

[2]

Proof:

To apply GMM on (3.4) and (3.7), the following quadratic formula must be minimized:

$$Q^+ = M^{+'} A^+ M^+ \quad (3.30)$$

Where A^+ is a $P^+ \times P^+$ positive definite weighting matrix.

Using (3.7) and (3.4) then substituting in (3.29) and (3.30), the quadratic formula in (3.30) can be rewritten as:

$$Q^+ = [(y^+ - Z^+ \delta)' W^+] A^+ [W^{+'} (y^+ - Z^+ \delta)] \quad (3.31)$$

Differentiating (3.31) with respect to δ and equating it with zero to obtain $\widehat{\delta}^+$ as:

$$\widehat{\delta}^+ = [Z^{+'} W^+ A^+ W^{+'} Z^+]^{-1} Z^{+'} W^+ A^+ W^{+'} y^+ \quad (3.32)$$

The weighting matrix in (3.30) might be taken as the inverse of the covariance matrix for moments conditions in (3.29):

$$A^+ = [W^{+'} \text{cov}(v^+) W^+]^{-1} \quad (3.33)$$

Where $\text{cov}(v^+)$ is a $2N(T-2) \times 2N(T-2)$ matrix that can be defined as:

$$\text{cov}(v^+) = \sigma_v^2 (I_N \otimes G^+) \quad (3.34)$$

where $G^+ = \begin{bmatrix} G & \mathbf{0}_{T-2} \\ \mathbf{0}_{T-2} & I_{T-2} \end{bmatrix}$ is $2(T-2) \times 2(T-2)$ matrix, I_N and I_{T-2} are identity matrices of order N and $(T-2)$ respectively, and $\mathbf{0}_{T-2}$ is a $(T-2) \times (T-2)$ matrix of zeros.

Using (3.34), the weighting matrix (A^+) in (3.33) can be rewritten as:

$$A^+ = \frac{1}{\sigma_v^2} [W^{+'} (I_N \otimes G^+) W^+]^{-1} \quad (3.35)$$

Using (3.35) and substituting in (3.32), the first step SYS GMM estimator can be defined as follows:

$$\widehat{\delta}_1^+ = \left[Z^{+'} W^+ (W^+ (I_N \otimes G^+) W^{+'})^{-1} W^{+'} Z^+ \right]^{-1} Z^{+'} W^+ (W^+ (I_N \otimes G^+) W^{+'})^{-1} W^{+'} y^+ \quad (3.36)$$

To find the second step SYS GMM estimator, the residuals of the first-step estimation \widehat{v}^+ is used to estimate the covariance matrix in (3.34) where $(W^{+'} (I_N \otimes G^+) W^+)$ in (3.36) is replaced with $V_N^+ = \sum_{i=1}^N W_i^{+'} v_i^+ v_i^{+'} W_i^+$ as follows:

$$\widehat{\delta}_2^+ = \left[Z^{+'} W^+ V_N^{-1} W^{+'} Z^+ \right]^{-1} Z^{+'} W^+ V_N^{-1} W^{+'} y^+ \quad (3.37)$$

where V_N^+ is $P^+ \times P^+$ matrix.

The covariance matrix of the second step SYS GMM estimator can be estimated as:

$$\widehat{\text{Cov}}(\widehat{\delta}_2^+) = \left[Z^{+'} W^+ V_N^{-1} W^{+'} Z^+ \right]^{-1} \quad (3.38)$$

3.3. SARGAN Overidentifying Restrictions Test

The model is considered overidentified when the number of moment conditions is greater than the number of parameters. In such case, a necessary test is needed to verify that orthogonality conditions are met. [1] proposed using Sargan test of overidentifying restrictions in case of DIF GMM estimation with the following null hypothesis and test statistic:

H_{0A} : The instruments are uncorrelated with the error terms and the model is correctly specified.

H_{1A} : The instruments are correlated with the error terms and the model is mis specified .

$$\mathcal{I} = \Delta \widetilde{v}' W \widetilde{V}_N^{-1} W' \Delta \widetilde{v} \sim \chi_{(P-k-1)}^2$$

Where $\widetilde{V}_N = \sum_{i=1}^N W_i' \Delta \widetilde{v}_i \Delta \widetilde{v}_i' W_i$, $\Delta \widetilde{v}$ are the differenced residuals resulting from the second step DIF GMM estimator, W is the matrix of instruments in (12), P is the number of columns of W .

Similarly, Sargan's test can be used in the case of SYS GMM under the following test statistic:

$$\mathcal{I}^+ = \widetilde{v}^{+'} W^+ \widetilde{V}_N^{+^{-1}} W^{+'} \widetilde{v}^+ \sim \chi_{(P^+-k-1)}^2$$

Where $\widetilde{V}_N^+ = \sum_{i=1}^N W_i^{+'} \widetilde{v}_i^+ \widetilde{v}_i^{+'} W_i^+$, \widetilde{v}^+ are the residuals resulting from the second step SYS GMM estimator, P^+ is the number of columns of W^+ .

4. Short Dynamic Panel Data Model With MIDAS

This section describes the approach given in [5] to estimate short dynamic panel data models in the presence of higher frequency regressors using DIF GMM MIDAS then extends it using SYS GMM MIDAS. The structure of the section is as follows: subsection 4.1 is dedicated to introducing the first difference GMM with MIDAS, then SYS GMM with MIDAS is illustrated in subsection 4.2, next modified Sargan test of overidentifying restriction is represented in subsection 4.3, finally inference about short dynamic panel data models with MIDAS is explained in subsection 4.4.

Short dynamic panel data model with MIDAS was introduced by [5] and can be written as follows:

$$y_{it} = \lambda y_{i,t-1} + x_{it}(\theta)\beta + u_{it} \quad (4.1)$$

$$x_{it}(\theta) = \sum_{g=1}^m x_{itg} \Pi_g(\theta) \quad (4.2)$$

$$u_{it} = \mu_i + v_{it} \quad (4.3)$$

Where $x_{it}(\theta)$ is $1 \times K$ vector of K weighted observations for the explanatory variables for individual (i) and time (t), $t = 3, \dots, T$ represents the lower frequency time dimension, $g = 1, \dots, m$ represents the higher frequency time dimension, x_{itg} are $1 \times k$ vectors of higher frequency regressors observations such that for each time point (t) there are m vectors for the regressor (i.e., x_{it1}, \dots, x_{itm}) corresponding to the observation of the dependent variable (y_{it}), $\Pi_g(\theta)$ is a weighting function used to parameterize the higher frequency regressor observations (x_{itg}) to be at the same frequency as the dependent variable (y_{it}). In this study the exponential Almon lag function with two parameters $\theta = (\theta_1, \theta_2)$ was used, that can be expressed as:

$$\Pi_g(\theta) = \frac{e^{g\theta_1 + g^2\theta_2}}{\sum_{g=1}^m e^{g\theta_1 + g^2\theta_2}}, \quad |\theta_1|, |\theta_2| \leq 1 \quad (4.4)$$

Under the following assumptions:

B1. $|\lambda| < 1$.

B2. $\mu_i \sim \text{IID}(0, \sigma_\mu^2)$.

B3. $v_{it} \sim \text{IID}(0, \sigma_v^2)$.

B4. θ is settled to fixed values $\theta_o = (\theta_{0,1}, \theta_{0,2})$.

B5. $E[\mu_i x_{it}(\theta_o)] \neq 0$.

B6. $E[v_{it} \mu_i] = 0$

B7. $E[v_{it} y_{it-j}] = 0; j = 1, \dots, t-1$.

B8. $x_{it}(\theta_o)$ is partitioned to $(x_{it}(\theta_o)^{(1)}, x_{it}(\theta_o)^{(2)})$ such that $x_{it}(\theta_o)^{(1)}$ is $1 \times K_1$ vector of strictly exogenous series $E[x_{it}(\theta_o)^{(1)'} v_{is}] = 0_{k_1} \quad 0t, s; s = 1, \dots, T$ and $x_{it}(\theta_o)^{(2)}$ is $1 \times K_2$ vector of predetermined series satisfying $E[x_{it}(\theta_o)^{(2)'} v_{is}] = 0_{k_2}, t \leq s$.

By taking the first difference for (4.1), the model can be rewritten as follows:

$$\Delta y_{it} = \lambda \Delta y_{i,t-1} + \Delta x_{it}(\theta) \beta + \Delta v_{it} \quad (4.5)$$

The model in (4.1) can be rewritten in matrix form as:

$$y = \lambda y_{-1} + X(\theta) \beta + u \quad (4.6)$$

where $X(\theta)$ is $N(T-2) \times K$ matrix of weighted series of K explanatory variables to be at the same frequency of the dependent variable.

Furthermore (4.5) can be rewritten as:

$$y = Z(\theta) \delta + u \quad (4.7)$$

where $Z(\theta) = [y_{-1} \quad X(\theta)]$ is $N(T-2) \times (K+1)$ matrix.

The differenced model in (4.5) might be written in matrix form for as follows:

$$\Delta y = \lambda \Delta y_{-1} + \Delta X(\theta) \beta + \Delta v \quad (4.8)$$

Moreover (4.7) can be rewritten as:

$$\Delta y = \Delta Z(\theta) \delta + \Delta v \quad (4.9)$$

Where $\Delta Z(\theta) = [\Delta y_{-1} \quad \Delta X(\theta)]$ is $N(T-2) \times (K+1)$ matrix.

4.1. First Difference GMM with MIDAS

In DIF GMM MIDAS, the matrix of valid instruments for each cross-sectional unit i depends on the choice of the fixed values of the weighting parameters vector θ_o and can be partitioned into three submatrices as follows:

$$W_i(\theta_o) = [W_i(y), W_i(x(\theta_o)^{(1)}), W_i(x(\theta_o)^{(2)})]$$

Where $W_i(y)$ is the same as defined in (3.9), $W_i(x(\theta_o)^{(1)}) = \text{diag}(x_{i1}(\theta_o)^{(1)}, \dots, x_{it}(\theta_o)^{(1)})$ is $(T-2) \times (T-2)(T+3)K_1/2$ matrix, $W_i(x(\theta_o)^{(2)}) = \text{diag}(x_{i1}(\theta_o)^{(2)}, \dots, x_{it-1}(\theta_o)^{(2)})$ is $(T-2) \times (T-2)(T+1)K_2/2$ matrix, $t = 3, \dots, T$.

Both $W_i(x(\theta_o)^{(1)})$ and $W_i(x(\theta_o)^{(2)})$ are similar to $W_i(x^{(1)})$ in (3.10) and $W_i(x^{(2)})$ in (3.11) respectively however here the explanatory variables series x_{itg} are weighted to be at the same frequency as the dependent variable series then the weighted series $x_{it}(\theta)$ in (4.2) are used as instruments.

For all cross-sectional units, the matrix of valid instruments in DIF GMM MIDAS is:

$$W(\theta_o) = [W_1(\theta_o)', \dots, W_N(\theta_o)'] \quad (4.10)$$

Where $W(\theta_o)$ is $N(T-2) \times P$ matrix.

The orthogonality conditions for the instrumental variables in the DIF GMM MIDAS method are:

$$E[y_{i,t-s} \Delta v_{it}(\theta_o)] = 0, \quad t = 3, \dots, T, \quad s = 2, \dots, T \quad (4.11)$$

$$E[x_{is}(\theta_o)^{(1)'} \Delta v_{it}(\theta_o)] = 0, \quad t = 3, \dots, T, \quad s = 1, \dots, T \quad (4.12)$$

$$E[x_{is}(\theta_o)^{(2)'} \Delta v_{it}(\theta_o)] = 0, \quad t = 3, \dots, T, \quad s = 1, \dots, T-1 \quad (4.13)$$

The orthogonality conditions in (4.9), (4.12) and (4.13) can be expressed using the matrix of valid instruments in (4.10) as:

$$M(\theta_o) = W(\theta_o)' \Delta v(\theta_o) \quad (4.14)$$

Where $M(\theta_o)$ is $P \times 1$ vector of orthogonality conditions.

Lemma 4.1: The first step DIF GMM MIDAS for the differenced model in (4.9) is:

$$\widehat{\delta}_1(\theta_o) = [\Delta Z(\theta_o)' W(\theta_o) (W(\theta_o)' (I_N \otimes G) W(\theta_o))^{-1} W(\theta_o)' \Delta Z(\theta_o)]^{-1} \times$$

$$[\Delta Z(\theta_o)' W(\theta_o) (W(\theta_o)' (I_N \otimes G) W(\theta_o))^{-1} W(\theta_o)' \Delta y]$$

Proof:

To apply GMM on (4.9), the following quadratic formula $Q(\theta_o)$ has to be minimized:

$$Q(\theta_o) = \mathbf{M}(\theta_o)' \mathbf{A}(\theta_o) \mathbf{M}(\theta_o) \quad (4.15)$$

Where $\mathbf{A}(\theta_o)$ is $P \times P$ positive definite weighting matrix.

Using (4.9) and substituting in (4.14) and (4.15), the quadratic formula in (4.15) can be rewritten as:

$$Q(\theta_o) = [(\Delta y - \Delta Z(\theta_o) \delta(\theta_o))' \mathbf{W}(\theta_o)] \mathbf{A}(\theta_o) [\mathbf{W}(\theta_o)' (\Delta y - \Delta Z(\theta_o) \delta(\theta_o))] \quad (4.16)$$

Differentiating (4.16) with respect to $\delta(\theta_o)$ and equating it with zero to obtain $\widehat{\delta}(\theta_o)$ as:

$$\widehat{\delta}(\theta_o) = [\Delta Z(\theta_o)' \mathbf{W}(\theta_o) \mathbf{A}(\theta_o) \mathbf{W}(\theta_o)' \Delta Z(\theta_o)]^{-1} \Delta Z(\theta_o)' \mathbf{W}(\theta_o) \mathbf{A}(\theta_o) \mathbf{W}(\theta_o)' \Delta y \quad (4.17)$$

The weighting matrix in (4.15) might be taken as the inverse of the covariance matrix for moments conditions in (4.14) as follows:

$$\mathbf{A}(\theta_o) = [\mathbf{W}(\theta_o)' \text{cov}(\Delta v(\theta_o)) \mathbf{W}(\theta_o)]^{-1} \quad (4.18)$$

Where $\text{cov}(\Delta v(\theta_o))$ is a $N(T-2) \times N(T-2)$ matrix that can be defined as:

$$\text{cov}(\Delta v(\theta_o)) = \sigma_{v(\theta_o)}^2 (\mathbf{I}_N \otimes \mathbf{G}) \quad (4.19)$$

Using (4.19), the weighting matrix $\mathbf{A}(\theta_o)$ in (4.18) can be rewritten as:

$$\mathbf{A}(\theta_o) = \frac{1}{\sigma_{v(\theta_o)}^2} [\mathbf{W}(\theta_o)' (\mathbf{I}_N \otimes \mathbf{G}) \mathbf{W}(\theta_o)]^{-1} \quad (4.20)$$

Using (4.20) and substituting in (4.17), the first step DIF GMM MIDAS might be expressed as follows:

$$\begin{aligned} \widehat{\delta}_1(\theta_o) &= \left[\Delta Z(\theta_o)' \mathbf{W}(\theta_o) (\mathbf{W}(\theta_o)' (\mathbf{I}_N \otimes \mathbf{G}) \mathbf{W}(\theta_o))^{-1} \mathbf{W}(\theta_o)' \Delta Z(\theta_o) \right]^{-1} \\ &\times \left[\Delta Z(\theta_o)' \mathbf{W}(\theta_o) (\mathbf{W}(\theta_o)' (\mathbf{I}_N \otimes \mathbf{G}) \mathbf{W}(\theta_o))^{-1} \mathbf{W}(\theta_o)' \Delta y \right] \end{aligned} \quad (4.21)$$

To get the second step DIF GMM MIDAS, the residuals of the first-step estimation $\widehat{\Delta v}(\theta_o)$ is used to estimate the covariance matrix in (4.19) where $(\mathbf{W}(\theta_o)' (\mathbf{I}_N \otimes \mathbf{G}) \mathbf{W}(\theta_o))$ in (4.21) is replaced with

$\mathbf{V}_N(\theta_o) = \sum_{i=1}^N \mathbf{W}_i(\theta_o)' \widehat{\Delta v}_i(\theta_o) \widehat{\Delta v}_i(\theta_o)' \mathbf{W}_i(\theta_o)$ as follows:

$$\widehat{\delta}_2(\theta_o) = \left[\Delta Z(\theta_o)' \mathbf{W}(\theta_o) \mathbf{V}_N(\theta_o)^{-1} \mathbf{W}(\theta_o)' \Delta Z(\theta_o) \right]^{-1} \left[\Delta Z(\theta_o)' \mathbf{W}(\theta_o) \mathbf{V}_N(\theta_o)^{-1} \mathbf{W}(\theta_o)' \Delta y \right] \quad (4.22)$$

where $\mathbf{V}_N(\theta_o)$ is $P \times P$ matrix.

The covariance matrix of the second step DIF GMM MIDAS estimator is as follows:

$$\text{Cov}(\widehat{\delta}_2(\theta_o)) = \left[\Delta Z(\theta_o)' \mathbf{W}(\theta_o) \mathbf{V}_N(\theta_o)^{-1} \mathbf{W}(\theta_o)' \Delta Z(\theta_o) \right]^{-1} \quad (4.23)$$

Special case:

Substituting in (4.21), (4.22), and (4.23) with the higher frequency time dimension $g = 1$, we get the first step DIF GMM, second step DIF GMM, covariance matrix of the second step DIF GMM defined in (3.23), (3.24), and (3.25) respectively.

4.2. System GMM with MIDAS

Short dynamic panel data model MIDAS approach suggested by [5] used DIF GMM, however this study extends the same approach using SYS GMM based on two simultaneous equations the first is the model based on levels in (4.1) and the second is the model based on first differences in (4.5).

Besides the orthogonality conditions defined in (4.11), (4.12), and (4.13) SYS GMM MIDAS can use the following additional orthogonality conditions:

$$E [\Delta y_{i,t-1} u_{it}(\theta_o)] = 0, \quad t = 3, \dots, T \tag{4.24}$$

$$E [\Delta x_{i,t-1} u_{it}(\theta_o)] = 0, \quad t = 3, \dots, T \tag{4.25}$$

In SYS GMM MIDAS, the matrix of valid instruments for each cross-sectional unit i can be defined as:

$$W_i(\theta_o)^+ = \begin{bmatrix} W_i(\theta_o) & 0 & 0_k & 0 & 0_k & \dots & 0 & 0_k \\ 0_P & [\Delta y_{i2} & \Delta x_{i2}(\theta_o)] & 0 & 0_k & \dots & 0 & 0_k \\ 0_P & 0 & 0_k & [\Delta y_{i3} & \Delta x_{i3}(\theta_o)] & \dots & 0 & 0_k \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_P & 0 & 0_k & 0 & 0_k & \dots & [\Delta y_{i,t-1} & \Delta x_{i,t-1}(\theta_o)] \end{bmatrix}$$

Where 0_P and 0_k are $1 \times P$ and $1 \times K$ vectors of zeros respectively, $W_i(\theta_o)^+$ is $2(T - 2) \times P^+$ matrix, and $P^+ = (T - 2) [(T + 1) + (T + 5) K] / 2$.

The vector of differenced idiosyncratic error terms $\Delta v(\theta_o)$ in (4.8) and the last $(T - 2)$ elements in the vector of error terms u in (4.7) can be stacked together as follows:

$$v_i(\theta_o)^+ = [\Delta v_{i3}(\theta_o), \dots, \Delta v_{it}(\theta_o), u_{i3}(\theta_o), \dots, u_{it}(\theta_o)]'$$

Where $v_i(\theta_o)^+$ is a $2(T - 2) \times 1$ vector.

For all cross-sectional units, the matrix of valid instruments in SYS GMM MIDAS is:

$$W(\theta_o)^+ = [W_1(\theta_o)^+, \dots, W_N(\theta_o)^+]' \tag{4.26}$$

Where $W(\theta_o)^+$ is $2N(T - 2) \times P^+$ matrix.

Similarly, $v_i(\theta_o)^+$ can be defined for all cross-sectional units as:

$$v(\theta_o)^+ = [v_1(\theta_o)^+, \dots, v_N(\theta_o)^+]'$$

Where $v(\theta_o)^+$ is $2N(T - 2) \times 1$ vector.

The orthogonality conditions for SYS GMM MIDAS can be expressed using the matrix of valid instruments in (4.26) as follows:

$$M(\theta_o)^+ = W(\theta_o)^+ v(\theta_o)^+ \tag{4.27}$$

Where $M(\theta_o)^+$ is $P^+ \times 1$ vector.

The vector of differenced observations for the dependent variable Δy in (4.9) and the vector of observations for the dependent variable y in (4.7) can be stacked together as follows:

$$y^+ = [\Delta y' \quad y']'$$

Where y^+ is $2N(T - 2) \times 1$ vector.

The matrix of differenced observations $\Delta Z(\theta_o)$ in (4.9) and the matrix of observations $Z(\theta_o)$ in (4.7) can be stacked together as follows:

$$Z(\theta_o)^+ = [\Delta Z(\theta_o)' \quad Z(\theta_o)']$$

Where $Z(\theta_o)^+$ is $2N(T - 2) \times (K + 1)$ matrix.

Lemma 4.2: The first step SYS GMM MIDAS of the model in (4.7) and the differenced model in (4.9) is:

$$\begin{aligned} \widehat{\delta}_1(\theta_o)^+ &= \left[Z(\theta_o)^+' W(\theta_o)^+ \left(W(\theta_o)^+ \left(I_N \otimes G^+ \right) W(\theta_o)^+ \right)^{-1} W(\theta_o)^+' Z(\theta_o)^+ \right]^{-1} \times \\ &\quad \left[Z(\theta_o)^+' W(\theta_o)^+ \left(W(\theta_o)^+ \left(I_N \otimes G^+ \right) W(\theta_o)^+ \right)^{-1} W(\theta_o)^+' y^+ \right] \end{aligned}$$

Proof:

To apply GMM on (4.7) and (4.9), the following quadratic formula has to be minimized:

$$Q(\theta_o)^+ = M(\theta_o)^+' A(\theta_o)^+ M(\theta_o)^+ \quad (4.28)$$

Where $A(\theta_o)^+$ is a $P^+ \times P^+$ positive definite weighting matrix.

Using (4.9) and (4.7) then substituting in (4.27) and (4.28), the quadratic formula in (4.28) can be rewritten as:

$$Q(\theta_o)^+ = \left[(y^+ - Z(\theta_o)^+ \delta(\theta_o))' W(\theta_o)^+ \right] A(\theta_o)^+ \left[W(\theta_o)^+' (y^+ - Z(\theta_o)^+ \delta(\theta_o)) \right] \quad (4.29)$$

Differentiating (4.29) with respect to δ and equating it with zero to obtain $\widehat{\delta}(\theta_o)^+$ as:

$$\widehat{\delta}(\theta_o)^+ = \left[Z(\theta_o)^+' W(\theta_o)^+ A(\theta_o)^+ W(\theta_o)^+' Z(\theta_o)^+ \right]^{-1} Z(\theta_o)^+' W(\theta_o)^+ A(\theta_o)^+ W(\theta_o)^+' y^+ \quad (4.30)$$

The weighting matrix in (4.28) might be taken as the inverse of the covariance matrix for moments conditions in (4.27) as follows:

$$A(\theta_o)^+ = \left[W(\theta_o)^+' \text{cov}(v(\theta_o)^+) W(\theta_o)^+ \right]^{-1} \quad (4.31)$$

Where $\text{cov}(v(\theta_o)^+)$ is a $2N(T - 2) \times 2N(T - 2)$ matrix that can be defined as:

$$\text{cov}(v(\theta_o)^+) = \sigma_{v(\theta_o)}^2 \left(I_N \otimes G^+ \right) \quad (4.32)$$

Using (4.32), the weighting matrix $A(\theta_o)^+$ in (4.31) can be rewritten as:

$$A(\theta_o)^+ = \frac{1}{\sigma_{v(\theta_o)}^2} \left[W(\theta_o)^+' \left(I_N \otimes G^+ \right) W(\theta_o)^+ \right]^{-1} \quad (4.33)$$

Using (4.33) and substituting in (4.30), the first step SYS GMM MIDAS can be defined as follows:

$$\begin{aligned} \widehat{\delta}_1(\theta_o)^+ &= \left[Z(\theta_o)^+' W(\theta_o)^+ \left(W(\theta_o)^+ \left(I_N \otimes G^+ \right) W(\theta_o)^+ \right)^{-1} W(\theta_o)^+' Z(\theta_o)^+ \right]^{-1} \times \\ &\quad \left[Z(\theta_o)^+' W(\theta_o)^+ \left(W(\theta_o)^+ \left(I_N \otimes G^+ \right) W(\theta_o)^+ \right)^{-1} W(\theta_o)^+' y^+ \right] \end{aligned} \quad (4.34)$$

To find the second step SYS GMM MIDAS, the residuals of the first-step estimation $\widehat{v}(\theta_o)^+$ is used to estimate the covariance matrix in (4.32) where $(W(\theta_o)^+ (I_N \otimes G^+) W(\theta_o)^+)$ in (4.34) is replaced with $V_N(\theta_o)^+ = \sum_{i=1}^N W_i(\theta_o)^+ v_i(\theta_o)^+ v_i(\theta_o)^+ W_i(\theta_o)^+$ as follows:

$$\widehat{\delta}_2(\theta_o)^+ = [Z(\theta_o)^+ W(\theta_o)^+ V_N(\theta_o)^{-1} W(\theta_o)^+ Z(\theta_o)^+]^{-1} \times [Z(\theta_o)^+ W(\theta_o)^+ V_N(\theta_o)^{-1} W(\theta_o)^+ y^+] \quad (4.35)$$

Where $V_N(\theta_o)^+$ is $P^+ \times P^+$ matrix.

The covariance matrix of the second step SYS GMM MIDAS can be estimated as:

$$\text{Cov}(\widehat{\delta}_2(\theta_o)^+) = [Z(\theta_o)^+ W(\theta_o)^+ V_N(\theta_o)^{-1} W(\theta_o)^+ Z(\theta_o)^+]^{-1} \quad (4.36)$$

Special case:

Substituting in (4.34), (4.35), and (4.36) with the higher frequency time dimension $g=1$, we get the first step SYS GMM, second step SYS GMM, covariance matrix of the second step SYS GMM defined in (3.36), (3.37), and (3.38) respectively.

Lemma 4.3: Assuming the dependent variable can be expressed by the model in (4.1), (4.2) and (4.2) while the higher frequency regressor x_{itg} is represented by panel AR (1) model of the lower frequency time dimension t as follows:

$$x_{itg} = \rho x_{i,t-1,g} + \eta_i + \epsilon_{itg} \quad (4.37)$$

Where $t = 2, \dots, T$, $i = 1, \dots, N$, $|\rho| < 1$ is a scalar, x_{itg} are $1 \times k$ vectors of higher frequency regressors observations, $x_{i,t-1,g}$ are $1 \times k$ vectors of higher frequency regressors lagged observations however the lag is in the lower frequency time dimension, η_i is fixed effects, ϵ_{itg} are error terms such that $\epsilon_{itg} \sim \text{IID}(0, \sigma_\epsilon^2)$ and independent of u_{it} in (4.1).

The initial conditions $y_{i,1}$ and $x_{i,1,g}$ that satisfies mean stationarity for the SYS GMM MIDAS are as follows:

$$x_{i,1,g} = \frac{\eta_i}{1-\rho} + e_{i,1,g}$$

$$y_{i,1} = \frac{\mu_i(1-\rho) + \beta\eta_i}{(1-\rho)(1-\lambda)} + w_{i,1}$$

Proof:

Expressing y_{it} in (4.1) recursively as follows:

When $t=2$:

$$y_{i2} = \lambda y_{i1} + x_{i2}(\theta)\beta + \mu_i + v_{i2}$$

When $t=3$:

$$y_{i3} = \lambda^2 y_{i1} + [\lambda x_{i2}(\theta) + x_{i3}(\theta)]\beta + [\lambda v_{i2} + v_{i3}] + [\lambda \mu_i + \mu_i]$$

When $t=4$:

$$y_{i4} = \lambda^3 y_{i1} + [\lambda^2 x_{i2}(\theta) + \lambda x_{i3}(\theta) + x_{i4}(\theta)]\beta + [\lambda^2 v_{i2} + \lambda v_{i3} + v_{i4}] + [\lambda^2 \mu_i + \lambda \mu_i + \mu_i]$$

In general:

$$y_{it} = \lambda^{t-1} \left(y_{i1} - \frac{\mu_i}{1-\lambda} \right) + \beta \sum_{s=0}^{t-2} \lambda^s x_{i,t-s}(\theta) + \sum_{s=0}^{t-2} \lambda^s v_{i,t-s} + \frac{\mu_i}{1-\lambda} \quad (4.38)$$

Similarly, x_{itg} in (4.37) can be expressed recursively as follows:

When $t=2$:

$$x_{i2g} = \rho x_{i1g} + \eta_i + \epsilon_{i2g}$$

When $t=3$:

$$x_{i3g} = \rho^2 x_{i1g} + [\rho \eta_i + \eta_i] + [\rho \epsilon_{i2g} + \epsilon_{i3g}]$$

When $t=4$:

$$x_{i4g} = \rho^3 x_{i1g} + [\rho^2 \eta_i + \rho \eta_i + \eta_i] + [\rho^2 \epsilon_{i2g} + \rho \epsilon_{i3g} + \epsilon_{i4g}]$$

In general:

$$x_{itg} = \rho^{t-1} \left[x_{i1g} - \frac{\eta_i}{1-\rho} \right] + \frac{\eta_i}{1-\rho} + \sum_{k=0}^{t-2} \rho^k \epsilon_{i,t-k,g} \quad (4.39)$$

Using (4.39), the weighted explanatory variables series $x_{it}(\theta)$ in (4.2) can be expressed as follows:

$$x_{it}(\theta) = \rho^{t-1} \left(x_{i1}(\theta) - \frac{\eta_i}{1-\rho} \right) + \frac{\eta_i}{1-\rho} + \sum_{k=0}^{t-2} \rho^k \epsilon_{i,t-k,g} \quad (4.40)$$

Where $x_{i1}(\theta) = \sum_{g=1}^m \Pi_g(\theta) x_{i1g}$.

Using (4.40), y_{it} in (4.1) can be represented as:

$$y_{it} = \lambda^{t-1} \left[y_{i1} - \frac{\mu_i(1-\rho) + \beta \eta_i}{(1-\rho)(1-\lambda)} \right] + \frac{\mu_i(1-\rho) + \beta \eta_i}{(1-\rho)(1-\lambda)} + \beta \sum_{s=0}^{t-2} \lambda^s \rho^{t-s-1} \left[x_{i1}(\theta) - \frac{\eta_i}{1-\rho} \right] + \beta \sum_{s=0}^{t-2} \lambda^s \sum_{k=0}^{t-s-2} \rho^k \epsilon_{i,t-s-k,g} + \sum_{s=0}^{t-1} \lambda^s v_{i,t-s} \quad (4.41)$$

From (4.39) and (4.41), the terms within the square brackets are the deviations of x_{itg} and y_{it} from their steady state paths, therefore mean stationarity can be assured by defining the initial conditions x_{i1g} and y_{i1} as:

$$x_{i,1,g} = \frac{\eta_i}{1-\rho} + e_{i,1,g} \quad (4.42)$$

$$y_{i,1} = \frac{\mu_i(1-\rho) + \beta \eta_i}{(1-\rho)(1-\lambda)} + w_{i,1} \quad (4.43)$$

4.3. Modified SARGAN Test of Overidentifying Restrictions

For short dynamic panel data models with MIDAS, [5] referred to Sargan test (subsection 3.3) as modified Sargan test of overidentifying restrictions.

In case of DIF GMM MIDAS, modified Sargan test statistic can be expressed as:

$$I(\theta_o) = \Delta \tilde{v}(\theta_o)' W(\theta_o) V_N(\theta_o)^{-1} W(\theta_o)' \Delta \tilde{v}(\theta_o) \sim \chi_{(P-k-1)}^2 \quad (4.44)$$

Where $V_N(\theta_o) = \sum_{i=1}^N W_i'(\theta_o) \Delta \tilde{v}_i(\theta_o) \Delta \tilde{v}_i'(\theta_o) W_i(\theta_o)$, and $\Delta \tilde{v}(\theta_o)$ are the differenced residuals resulting from the second step DIF GMM MIDAS.

Sargan's test of overidentifying restrictions for SYS GMM MIDAS can be expressed as:

$$I(\theta_o)^+ = \tilde{v}(\theta_o)^+{}' W(\theta_o)^+ V_N(\theta_o)^{+^{-1}} W(\theta_o)^+ \tilde{v}(\theta_o)^+ \sim \chi_{(p-k-1)}^2 \quad (4.45)$$

Where $V_N(\theta_o)^+ = \sum_{i=1}^N W_i^{+'}(\theta_o) \Delta \tilde{v}_i^+(\theta_o) \Delta \tilde{v}_i^{+'}(\theta_o) W_i^+(\theta_o)$, and $\tilde{v}(\theta_o)^+$ are the residuals resulting from the second step SYS GMM MIDAS.

4.4. Inference About Short Dynamic Panel Data Models with MIDAS

Under the assumptions from B1 to B8, inference about the parameters of model (4.7) or (4.9) with null hypothesis $H_{OB}: \delta_r = \delta_{ro}$ can be carried out using the following t-test statistic:

$$t = \frac{\widehat{\delta}_r - \delta_{ro}}{\widehat{s.e.}(\widehat{\delta}_r)} \quad (4.46)$$

Where $r = 1, \dots, (K+1)$ such that $(K+1)$ is the number of parameters, $\widehat{\delta}_r$ and $\widehat{s.e.}(\widehat{\delta}_r)$ are the corresponding elements of the second step DIF GMM MIDAS and the square root of the elements on the main diagonal for the covariance matrix of the second step DIF GMM MIDAS given by (4.22) and (4.23) respectively or the corresponding elements of the second step SYS GMM MIDAS and the square root of the elements on the main diagonal for the covariance matrix of the second step SYS GMM MIDAS given by (4.35) and (4.36) respectively, and t follows asymptotically standard normal distribution.

Inference about the nuisance parameters vector θ in models (4.7) or (4.9) can be done by inverting the modified Sargan test statistic given by (4.44) in case of DIF GMM MIDAS or by (4.45) in case of SYS GMM MIDAS. The inversion can be carried out numerically using grid search to find the set of all values of θ retained by the test meanwhile an empty set indicates a lack of fit. In case of using two-parameters exponential Almon lag function for weighting, where $-1 \leq \kappa_1, \kappa_2 \leq 1$, [5] used a grid search that was constructed of an outer loop with increment 0.01 starting with -1 and ending with 1 which gives 201 options for the first parameter, and an inner loop with the same increment and the same number of options, making the number of alternatives for the grid search as $201^2 = 40401$ alternatives for $\theta = (\theta_1, \theta_2)$. For each alternative, the model is fitted and orthogonality conditions are tested using Sargan's test where the set of pairs for (θ_1, θ_2) that has p-value greater than 0.05 considered as confidence region for exponential Almon lag parameters, while the pair (θ_1, θ_2) and the values for $\widehat{\delta}_r$ that correspond the highest p-value (that is greater than 0.05) are considered as model estimate.

As noticed by [5], there are three alternatives to make inference about the model using grid search as described above where these methods rely on modified Sargan test statistic or the t-test statistic:

1. The first alternative is to test the joint null hypothesis $H_{OC}: \theta = \theta_o$ and $\delta = \delta_o$ relying on Sargan test statistic $I(\theta_o, \delta_o)$ where the null hypothesis is not rejected when $I(\theta_o, \delta_o) < \chi_{(p-k-1)}^2$. A joint confidence region $C(\theta, \delta; \alpha)$ can be constructed by considering all the values of θ and δ satisfying $I(\theta_o, \delta_o) < \chi_{(p-k-1)}^2$ which can be described as

$$C(\theta, \delta; \alpha) = \left\{ \theta_o, \delta_o; I(\theta_o, \delta_o) < \chi_{(\alpha, p-k-1)}^2 \right\}$$

or using the p-value of Sargan test as

$$C(\theta, \delta; \alpha) = \{\theta_o, \delta_o; p_{I(\theta_o, \delta_o)} > \alpha\}$$

An empty confidence region indicates a lack of fit while a point estimate is declared to be the one with the highest p-value of Sargan test.

1. The second alternative is finding a set of parameters θ_o satisfying a valid specification of the model $I(\theta_o, \delta_o) < \chi^2_{(p-k-1)}$ as a preliminary step where under the null hypothesis $H_{0D}: \theta = \theta_o$ with a level of significance α_1 .

$$C(\theta; \alpha_1) = \{\theta_o; I(\theta_o) < \chi^2_{(\alpha_1, (p-k-1))}\}$$

or

$$C(\theta; \alpha_1) = \{\theta_o; p_{I(\theta_o)} > \alpha_1\}$$

Then, to make inference about δ , the null hypothesis $H_o: \delta = \delta_o$ is defined with a conservative level of significance $\alpha = \alpha_1 + \alpha_2$, then a search over the retained weighting parameters set $C(\theta; \alpha_1)$ can be performed to find the smallest t-ratio value.

$$t_C = \inf_{\theta \in C(\theta; \alpha_1)} \{|t(\delta_o, \theta)|\}$$

Where inf is the minimum value of a given set.

The null hypothesis is not rejected when the value t_C is less than the critical value at a significance level α_2 i.e., $t_C < t_{crit(\alpha_2)}$. The confidence interval for δ can be constructed through the set of non-rejected values of δ by finding the smallest and largest values of the set.

$$C(\delta; \alpha) = \{\delta_o; t_C < t_{crit(\alpha_2)}\}$$

Where $t_{crit(\alpha_2)}$ is the corresponding standard score at level of significance α_2 .

1. The last alternative is to test a hypothesis about δ under the null hypothesis $H_{0E}: \delta = \delta_o$ with the level of significance α where the full MIDAS parameters set $\theta \in \Theta$ can be searched over to find the smallest t-ratio value.

$$t_A = \inf_{\theta \in \Theta} \{|t(\delta_o, \theta)|\}$$

Where Θ is the set of all alternatives of the weighting function parameters.

Similarly, the null hypothesis is not rejected when this value is less than the critical value at significance level α , i.e., $t_A < t_{crit(\alpha)}$. The confidence interval for δ can be constructed through the set of non-rejected values of δ_o by finding the smallest and largest values of the set.

$$C(\delta; \alpha) = \{\delta_o; t_A < t_{crit(\alpha)}\}$$

$t_{crit(\alpha)}$ is the corresponding standard score at level of significance α .

5. Simulation Study

In this section, the Monte Carlo simulation is used to calculate the power and empirical size of the t-test given by (4.46) under different alternatives for the weighting function parameters vector (θ) and the weighted higher frequency regressor coefficient (β) to compare between DIF GMM MIDAS and SYS GMM MIDAS.

5.1. Simulation Design

Data generating process of the simulations for the dependent variable are based on the model in (4.1), (4.2) and (4.3) while the data generating process for the higher frequency regressor are based on in (4.37). The choice of the model in (4.37) to be AR(1) of the lower frequency time dimension rather than being AR(1) of the higher frequency time dimension to avoid correlation between MIDAS weights and the remainder disturbance term as discussed in Khalaf et al. [5]. The initial conditions were set to satisfy mean stationarity (lemma 4.3) as provided in (4.42) and (4.43).

For DIF GMM MIDAS, the simulation design was like Khalaf et al. [5] where the lagged dependent variable coefficient λ in (39) was 0.5, the weighted higher frequency regressor coefficient β in (39) was 2, the values of the specific individual effects μ_i in (41) were generated from $N(0, 1)$, the values of the idiosyncratic error terms v_{it} in (41) were generated from $N(0, 1)$, the weighting function $\Pi_g(\theta)$ used to parametrize the higher frequency regressor in (40) was the two parameters exponential Almon lag function in (42) with $(\theta_1, \theta_2) = (0, 0.05)$, the coefficient of the lagged regressor ρ in (75) was 0.8, the fixed effects η_i in (75) were set to be 0, the values of the error term ϵ_{itg} in (75) were generated from $N(0, 0.9)$, the number of cross-sectional units N was 500, the number of lower frequency time points T was 5, the number of higher frequency time points m was 20, the number of replications for each experiment was 2000 with fixed seed each runtime, the set of chosen values for the exponential Almon lag parameters (θ_1, θ_2) as null hypothesis was $\{(0, 0.1), (0, 0.05), (0, 0), (0, -0.05), (0, -0.1)\}$ where the value of θ_1 was fixed to 0 since the power and empirical size are much more dependent on variations in θ_2 as pointed out in [5], the set of values for β as null hypothesis was $\{-1, -0.5, 0, 0.5, 1, 2, 3, 4\}$, the nominal size of the t-test statistic was 0.05, the power and empirical size can be estimated as follows:

$$\text{power} = \frac{a_1}{2000}$$

$$\text{empirical size} = \frac{a_2}{2000}$$

Where a_1 is the number of times the null hypothesis was rejected while it was different from the data generating process, and a_2 is the number of times the null hypothesis was rejected while it was the same as the data generating process.

As pointed out by [18], most of the Monte Carlo studies concerning SYS GMM estimators for the short dynamic panel data models rely on the mean stationarity as a sufficient condition for consistency by generating $T+S$ time observations then dropping S observations and using the rest of them where S is more than or equal 50. The simulation design for SYS GMM MIDAS was the same as DIF GMM MIDAS however the number of used time points $T+S$ was 55 such that the number of dropped time observations S was 50 while the number of lower frequency time points T was 5, additionally the values of the error term $e_{i,1,g}$ in (4.42) were generated from $N(0, 0.9)$ and the values of the error term $w_{i,1}$ in (4.43) were generated from $N(0, 1)$.

5.2. Simulation Results

Table 1 represents the Probability of rejecting the null hypothesis $H_0: (\theta_1, \theta_2) = (\theta_{01}, \theta_{02})$ and $\beta = \beta_0$ using t-test statistic given by (4.46) in case of DIF GMM MIDAS compared to SYS GMM MIDAS. There are 5 alternatives for the exponential Almon lag parameters $(\theta_{01}, \theta_{02})$ and 8 alternatives for the coefficient of weighted higher frequency regressor (β_o) with total of 40 alternatives as null hypothesis for model parameters. The t-test provides more power in the case of SYS GMM MIDAS than DIF GMM MIDAS while both have empirical size approximately equals to the nominal size. Most of the power values are higher than 0.6. Low power values can be highlighted and explained as follows:

1. : $H_0: (\theta_1, \theta_2) = (0, 0.05)$ and $\beta = 2$, the values of the empirical size of both estimators are approximately equal to the nominal size of $\alpha = 0.05$.
2. : $H_0: (\theta_1, \theta_2) = (0, 0)$ and $\beta = 2$, the value observed represents the power of the test in detecting the false null hypothesis. According to the values of both estimators, a higher power can be detected in SYS GMM MIDAS (0.925) compared to DIF GMM MIDAS (0.235). These results can be explained as a product of having $(\theta_1, \theta_2) = (0, 0)$; which is equivalent to the average of the higher frequency data.
3. : $H_0: (\theta_1, \theta_2) = (0, 0)$ and $\beta = 3$ or 4 , the values indicate the ability of detecting that the null hypothesis is different than the data generating process. Slightly higher power can be recognized for the SYS GMM MIDAS compared to DIF GMM MIDAS in case of $(\theta_1, \theta_2) = (0, 0)$ which is equivalent to using the average to weight the higher frequency regressor.
4. : $H_0: \beta$ is between -0.5 and 0.5 , small power values realized when $\beta = 0$ or closer to 0 where the MIDAS covariate $x_{it}(\theta)$ simply disappears from the model and excluded from the matrix of instruments given by (4.10) or (4.26). In case of DIF GMM MIDAS, the values listed as 0.195, 0.04, and 0.214 corresponding to $\beta = -0.05$, $\beta = 0$ and $\beta = 0.05$ respectively when $(\theta_1, \theta_2) = (0, -0.05)$. Also, the values 0.289, 0.038, and 0.298 corresponding to $\beta = -0.05$, $\beta = 0$ and $\beta = 0.05$ respectively when $(\theta_1, \theta_2) = (0, -0.1)$. In case of SYS GMM MIDAS, the value 0.058 corresponding to $\beta = 0$ when $(\theta_1, \theta_2) = (0, -0.05)$. Likewise, the value 0.06 corresponding to $\beta = 0$ when $(\theta_1, \theta_2) = (0, -0.1)$.

6. Concluding Remarks

This study is intended to examine the impact of including a higher frequency regressor(s) in the short dynamic panel data models. The difference and system GMM estimator extensively discussed in the balanced case and in the presence of higher frequency regressor(s). In the presence of a higher frequency regressor, by assuming AR (1) of the lower frequency time dimension as data generating process for the regressor(s) the initial conditions satisfying mean stationarity for SYS GMM MIDAS were derived. A simulation study was conducted to calculate the power and empirical size of t-test statistic given by (4.46) where SYS GMM MIDAS was found to provide more power than DIF GMM MIDAS, additionally the average weighting corresponding to exponential Almon lag function $(\theta_1, \theta_2) = (0, 0)$ gave low power values, similarly testing the significance of the weighted higher frequency regressor parameter $\beta = 0$ or close to zero had low power values.

Table 1. Probability of Rejecting the Joint Null Hypothesis $H_o: (\theta_1, \theta_2) = (\theta_{01}, \theta_{02})$ and $\beta = \beta_0$

H_o	θ_0		β_0							
	θ_{01}	θ_{02}	-1	-0.5	0	0.5	1	2	3	4
DIF GMM MIDAS	0	0.1	1	1	1	1	1	0.797	1	1
	0	0.05	1	1	1	1	1	0.049	1	1
	0	0	0.999	0.993	0.972	0.898	0.747	0.235	0.047	0.217
	0	-0.05	0.637	0.195	0.04	0.214	0.659	0.993	1	1
	0	-0.1	0.798	0.289	0.038	0.298	0.807	0.999	1	1
SYS GMM MIDAS	0	0.1	1	1	1	1	1	0.998	1	1
	0	0.05	1	1	1	1	1	0.052	1	1
	0	0	1	1	1	1	1	0.925	0.219	0.276
	0	-0.05	0.977	0.549	0.058	0.551	0.985	1	1	1
	0	-0.1	0.997	0.709	0.06	0.693	0.998	1	1	1

Conflict of interest

The authors declare no conflict of interests.

References

1. Arellano, M., and Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The review of economic studies*, 58(2), 277-297.
2. Blundell, R., and Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics*, 87(1), 115-143.
3. Ghysels, E., Santa-Clara, P., and Valkanov, R. (2004). The MIDAS Touch: Mixed Data Sampling Regression Models. (UCLA Working Paper).
4. Forni, C., Marcellino, M., and Schumacher, C. (2013). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 178(1), 57-82.
5. Khalaf, L., Kichian, M., Saunders, C. J., and Voia, M. (2021). Dynamic panels with MIDAS covariates: Nonlinearity, estimation and fit. *Journal of Econometrics*, 220(2), 589-605.
6. Anderson, T. W., and Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American statistical Association*, 76(375), 598-606.
7. Bhargava, A., and Sargan, J. D. (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica: Journal of the Econometric Society*, 1635-1659.
8. Keane, M. P., and Runkle, D. E. (1992). On the estimation of panel-data models with serial correlation when instruments are not strictly exogenous. *Journal of Business and Economic Statistics*, 10(1), 1-9.

9. Arellano, M., and Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of econometrics*, 68(1), 29-51.
10. Ahn, S. C., and Schmidt, P. (1995). Efficient estimation of models for dynamic panel data. *Journal of econometrics*, 68(1), 5-27.
11. Kiviet, J. F. (1995). On bias, inconsistency, and efficiency of various estimators in dynamic panel data models. *Journal of econometrics*, 68(1), 53-78.
12. Alonso-Borrego, C., and Arellano, M. (1999). Symmetrically normalized instrumental-variable estimation using panel data. *Journal of Business and Economic Statistics*, 17(1), 36-49.
13. Hsiao, C., Pesaran, M. H., and Tahmiscioglu, A. K. (2002). Maximum likelihood estimation of fixed effects dynamic panel data models covering short time periods. *Journal of econometrics*, 109(1), 107-150.
14. Hsiao, C., and Tahmiscioglu, A. K. (2008). Estimation of dynamic panel data models with both individual and time-specific effects. *Journal of Statistical Planning and Inference*, 138(9), 2698-2721.
15. Kruiniger, H. (2013). Quasi ML estimation of the panel AR (1) model with arbitrary initial conditions. *Journal of Econometrics*, 173(2), 175-188.
16. Bai, J., 2013, Fixed-effects dynamic panel data models, a factor analytical method, *Econometrica* 81, 285-314.
17. Moral-Benito, E. (2013). "Likelihood-Based Estimation of Dynamic Panels with Predetermined Regressors." *Journal of Business Economic Statistics* 31, 451–472.
18. Bun, M. J., and Sarafidis, V. (2015). Dynamic panel data models. In Baltagi, B. H. (Ed.), *The Oxford handbook of panel data* (pp. 76-110).
19. Bun, M. J., Carree, M. A., Juodis, A. (2017). On maximum likelihood estimation of dynamic panel data models. *Oxford Bulletin of Economics and Statistics*, 79(4), 463-494.
20. Bricongne, J. C., Meunier, B., and Pical, T. (2021). Can satellite data on air pollution predict industrial production. (BDF Working Paper No. 847). Retrieved from Banque de France website: <https://www.banque-france.fr/>
21. Alvarez, J., and Arellano, M. (2003). The time series and cross-section asymptotics of dynamic panel data estimators. *Econometrica*, 71(4), 1121-1159.
22. Almongy, H. M., & Almetwally, E. M. (2020). Robust estimation methods of generalized exponential distribution with outliers. *Pakistan Journal of Statistics and Operation Research*, 16(3), 545-559.
23. Almetwally, E. M., & Almongy, H. (2018). Comparison between m estimation, s estimation, and mm estimation methods of robust estimation with application and simulation. *International Journal of Mathematical Archive*, 9(11), 1-9.
24. Alshenawy, F. Y., & Almetwally, E. M. (2023). A comparative study of statistical and intelligent classification models for prediction diabetes. *Advances and Applications in Statistics*, 88(2), 201-223.
25. Kamel, A. R., & Abonazel, M. R. (2023). A Simple Introduction to Regression Modeling using R. *Computational Journal of Mathematical and Statistical Sciences*, 2(1), 52-79.

-
26. Moral-Benito, E., Allison, P., and Williams, R. (2019). Dynamic panel data modelling using maximum likelihood: an alternative to Arellano-Bond. *Applied Economics*, 51(20), 2221-2232.



© 2023 , licensee the scientific association for studies and applied research (SASAR). This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)