TF-MIDAS: a new mixed-frequency model to forecast macroeconomic variables

Bonino-Gayoso, Nicolás and García-Hiernaux, Alfredo

Universidad de la República, Universidad Complutense de Madrid

30 March 2019

TF-MIDAS: a new mixed-frequency model to forecast macroeconomic variables

Nicolás Bonino-Gayoso* and Alfredo Garcia-Hiernaux†

Preliminary version: March, 2019

Abstract

This paper tackles the mixed-frequency modeling problem from a new perspective. Instead of drawing upon the common distributed lag polynomial model, we use a transfer function representation to develop a new type of models, named TF-MIDAS. We derive the theoretical TF-MIDAS implied by the high-frequency VARMA family models and as a function of the aggregation scheme (flow and stock). This exact correspondence leads to potential gains in terms of nowcasting and forecasting performance against the current alternatives. A Monte Carlo simulation exercise confirms that TF-MIDAS beats UMIDAS models in terms of out-of-sample nowcasting performance for several data generating high-frequency processes.

Keywords: mixed-frequency models, TF-MIDAS, U-MIDAS, Nowcasting, Forecasting.
JEL: C22, C32, N70, F15

1 Introduction

Economic policy-makers, entrepreneurs and investors, among other agents, need to have access to real-time assessments of the state of the economy, along with nowcasts and forecasts of its expected evolution. The sooner they have access to information about the real economic situation, the better prepared they will be to make decisions to update their initial plans. This becomes even more important in non-stable economic environments.

Unfortunately, data offered by the System of National Accounts (SNA) is delivered with considerable delay. In the case of European countries, Eurostat provides the value of EU and euro area GDP 70 days after the end of quarter, preceded by a first preliminary

* PhD candidate at Universidad Complutense de Madrid; Decon, Universidad de la República, Uruguay. E-mail: nicolas.bonino.gayoso@gmail.com
† ICAE, Universidad Complutense de Madrid. E-mail: agarciah@ucm.es

estimate and a second estimate 30 and 45 days after the end of quarter, respectively.[1] This delay, as Castle and Hendry (2013) point out, is the consequence of the existence of several difficulties for producing timely and accurate low-frequency aggregates: (i) not all disaggregated data are available when is needed to compute a relevant aggregate; (ii) many disaggregated time series are only preliminary estimates, subject to substantial revisions, so they are not accurate descriptions of the current conditions.

On the other hand, there is a considerable number of short-term economic indicators available at a much earlier stage that could be used to extract information about the state of the economy. Thus we have access to monthly data from consumer surveys or the industrial production index, daily data from financial markets and even more frequently observed variables, such as Google or Twitter trends and mobile phone data. Although probably not as complete as SNA data, these indicators could anticipate relevant information.

Forecasting models usually require to use data observed at the same frequency, which represents a setback when using a more complete information set provided by mixed-frequency variables. Several solutions have been developed to try to overcome this problem. In the following lines we will present the most common ones.

Throughout the paper we follow the notation used in Foroni et al. (2015). A high-frequency indicator is denoted by letter x. Let t be the time index for this variable x, $t = 1, ..., T$ (e.g., although not exclusively, months), being T the last period for which data of variable x is available. The lag operator for this high-frequency indicator is denoted by L. If x_t is the monthly industrial production index used to nowcast quarterly GDP, then Lx_t is the value of this index corresponding to the previous month.

Let y be the low-frequency variable that is aimed to be nowcast, sampled at periods denoted by time index $t_q = 1, ..., T_q$ (e.g., although not exclusively, quarters), being T_q the last period for which data of variable y is available. Usually, $T \geq k * T_q$, as observations of high-frequency indicators are available earlier than low-frequency ones. Past realizations of the low-frequency variable will be denoted by the lag operator Z, where $Z = L^k$. So, if y_{t_q} is quarterly GDP, then Zy_{t_q} would be GDP of the previous quarter.

The high-frequency indicator x is sampled k times between samples of y. For example, for quarterly GDP and a monthly indicator, $k = 3$.

Finally, aggregated values of the high-frequency indicator are denoted by x_{t_q}, e.g., quarterly aggregation of the monthly industrial production index. Both, the target variable y and the indicator x are assumed to be stationary, so these variables often correspond to a (log) differenced version of some raw series z.

The way to extract information from the available indicators is not straightforward and there are several methodologies, with different levels of complexity, to address this task. Several classes of models have been proposed to work explicitly with mixed-frequency datasets, most widely used being MIDAS (MIxed DAta Sampling) family of models.

[1] In the case of USA, the Bureau of Economic Analysis (BEA) releases a first estimate of GDP 28 days after the end of quarter, followed by a second and a third estimate 60 and 90 days after the end of quarter, respectively. In Latin American countries the schedule is similar. In Uruguay, for example, GDP value is released by the Central Bank 75 days after the end of quarter.

MIDAS models (Ghysels et al., 2002, 2003, 2006) are defined in terms of a Distributed Lag (DL) polynomial, explicitly modelling the relationship between variables observed at different frequencies. In order to keep parsimony, standard MIDAS models are defined in terms of a short number of parameters. This kind of model has been applied to nowcasting GDP, private consumption and corporate bond spreads, among other variables (Ghysels et al., 2007; Clements and Galvão, 2008, 2009; Bai et al., 2013; Schumacher, 2014; Duarte et al., 2017).

A specific variation of the standard model, known as Unrestricted MIDAS (U-MIDAS), is introduced by Foroni et al. (2015). Based on a series of simulation exercises, these authors state that U-MIDAS' nowcasting accuracy outperforms that one of standard MIDAS, when the difference in sampling frequencies is not high, specially for monthly to quarterly frequency, as it is usually the case of macroeconomic nowcasting.

In this paper we apply a transfer function representation to the Distributed Lag (DL) polynomial, deriving a new type of mixed-frequency model, named Transfer Function-MIDAS or, simply, TF-MIDAS.

We consider that the representation of the classical MIDAS model in terms of ratios of finite lag polynomials is a more appropiate way to model mixed-frequency data, as it is a more general alternative than current DL models. Moreover, a formal methodology of specification and estimation was already developed (see, e.g, Box and Jenkins, 1976) and tested.

We present evidence from Monte Carlo simulation exercises confirming that TF-MIDAS beats currently available UMIDAS models in terms of out-of-sample nowcasting performance. Working with simulated data allows us to consider different variants of Data Generating Processes (DGP) and so to identify for which specific DGPs the advantage of TF-MIDAS is significantly greater. Following Foroni et al. (2015), we initially consider a HF-VAR(1) process, and we find that there is no considerable difference in nowcasting performance between U-MIDAS and TF-MIDAS, whether we consider a stock or a flow variable.

We also consider the cases of a HF-VMA(1) and HF-VMA(3) processes, where we find that TF-MIDAS shows a considerable better performance than U-MIDAS. In the case of a HF-VMA(1) DGP, no remarkable advantage is observed in the case of a stock variable, but the results change completely when a flow variable is considered, as we find a significant advantage of TF-MIDAS relative to UMIDAS. In the case of a HF-VMA(1) process, TF-MIDAS proves to have a better nowcasting performance whether we consider a stock or a flow variable.

We conclude that TF-MIDAS presents an overall much better nowcating performance than U-MIDAS. In the case of a HF-VAR (1) process the out-of-sample nowcasting performance of both models is practically the same, whereas in the case of a VMA(1) and VMA(3) processes TF-MIDAS outstands as a more accurate nowcasting model. These results are robust to different parameter specifications.

The paper is organized in five sections, including the present introduction. In section 2, MIDAS models are briefly reviewed, specially focusing on one specific variant of these models known as U-MIDAS (Unrestricted MIDAS). Section 3 is dedicated to introduce our new proposed model, TF-MIDAS. Theoretical elements are presented and empirical issues are discussed in this section. In section 4, the simulation design and the models under comparison are detailed and results of relative out-of-sample nowcasting performance of TF-MIDAS model are presented and discussed. We determine for which kind of processes TF-MIDAS shows a significantly better relative performance and so it would be advisable to apply it. Finally, in section 5 we summarize the main results and conclude.

2 Review of MIDAS models

Prior to review the specific U-MIDAS model, it is worth presenting a brief summary of the original MIDAS (MIxed DAta Sampling) model, which was introduced for the first time in two working papers (Ghysels et al., 2002, 2003) and later in a published paper (Ghysels et al., 2006).

In MIDAS models the response of the low-frequency variable to a high-frequency explicative variable is modelled through a distributed lag polynomial and special attention is paid to parsimony. In order to avoid the so-called "parameter proliferation" problem, lag coefficients are not free, but are defined as a function of a vector of few parameters (θ).

Ghysels et al. (2002) propose the following DL specification for a MIDAS model with N explicative variables:

$$\omega(L)y_t \;=\; \beta_0 \;+\; \sum_{i=1}^{N} \beta_i B_i(L;\theta)x_{i,t-1} \;+\; \epsilon_t \qquad\qquad t = k, 2k, ..., Tk, \tag{1}$$

where $B_i(L;\theta) = \sum_{j=0}^{K} b_i(j;\theta)L^j$ and $\omega(L) = \omega_0 + \omega_1 L + ... + \omega_{k-1}L^{k-1}$ determines the aggregation scheme.[2]

Function $b(j,\theta)$, component of the lag polynomial, is used to model the weights assigned to each lag of the high-frequency indicator. This function depends on the indicator's period, j, and a vector of hyperparameters, θ. An overview of different weighting functions proposed so far in the literature is provided in Ghysels (2014), the most popular being Exponential Almon and Beta.[3]

There are several variations constructed upon the basic MIDAS model. A summary of the main features of these variations can be found in Foroni and Marcellino (2013).

[2] In the case of a stock variable, it would be $\omega_0 = 1$, $\omega_1 = ... = \omega_{k-1} = 0$. In the case of a flow variable, it would be $\omega_0 = \omega_1 = ... = \omega_{k-1} = 1$, if the values are aggregated by addition, or $\omega_0 = \omega_1 = ... = \omega_{k-1} = 1/k$, if the values are aggregated by average.

[3] The Almon Exponential weighting function was proposed in Ghysels et al. (2005) and it has the following expression, with Q shape parameters:

$$b(j;\theta) = \frac{exp(\theta_1 j + ... + \theta_Q j^Q)}{\sum_{j=0}^{K} exp(\theta_1 j + ... + \theta_Q j^Q)} \tag{2}$$

Beta weighting function, proposed for the first time in Ghysels et al. (2003), includes only two shape parameters:

2.1 U-MIDAS model

Foroni et al. (2012, 2015) propose a variant of MIDAS model known as Unrestricted MIDAS (U-MIDAS), which does not employ functional distributed lag polynomials to model the relationship between x and y.

When the difference in sampling frequencies is not high (as it occurs when monthly indicators are used to nowcast quarterly GDP) the risk of falling into the curse of dimensionality becomes less relevant and so it does the need to resort to functional distributed lag polynomials.

The U-MIDAS model based on a linear lag polynomial is defined by:

$$C(L^k)\omega(L)y_t \;=\; \delta_1(L)x_{1,t-1} \;+ ... +\; \delta_N(L)x_{N,t-1} \;+\; \epsilon_t \qquad\qquad t = k, 2k, ..., Tk, \quad (4)$$

where $C(L^k) = 1 - c_1L^k - ... - c_cL^{kc}$; $\delta_j(L) = \delta_{j,0} + \delta_{j,1}L + ... + \delta_{j,v_j}L^{v_j}$, $j = 1, ..., N$; x_j, $j = 1, ..., N$ are the explanatory variables, each of them affecting y up to lag v_j. Foroni et al. (2015) state that notwithstanding the error term ϵ_t has in general a moving average structure, i.e., $\epsilon_t = C(L^k)a_t$, where $a_t \sim iid(0, \sigma^2)$, they suggest to consider an AR approximation for simplicity. Finally, $\omega(L) = \omega_0 + \omega_1L + ... + \omega_{k-1}L^{k-1}$ determines the aggregation scheme.

Foroni et al. (2015) state that basic MIDAS model can be thought of as nested in U-MIDAS specification, because it is the result of imposing a particular dynamic pattern on it. An important computational advantage of U-MIDAS model over the basic MIDAS model is that it can be estimated by simple OLS, as long as lag orders c and v_j are long enough to make the error term ϵ_t uncorrelated.

Once the model is fitted, nowcasting for y_{Tk+k} conditional on information available at period $Tk + k - 1$ would be expressed as:

$$\hat{y}_{Tk+k|Tk+k-1} \;=\; (\hat{c}_1L^k + ... + \hat{c}_cL^{kc})y_{Tk+k} \;+\; \hat{\delta}_1(L)x_{1,Tk+k-1} \;+\; ... \;+\; \hat{\delta}_N(L)x_{N,Tk+k-1}. \quad (5)$$

3 TF-MIDAS

In the literature about mixed-frequency models, MIDAS models are represented in terms of a DL polynomial expression, as they were shown in the previous section. In this section we present an alternative representation based on a Transfer Function (TF), which we refer to as TF-MIDAS. We also derive the new proposed model from a general linear dynamic model and compare it, from a theoretically viewpoint, with a U-MIDAS model.

$$b(j; \theta_1, \theta_2) = \frac{f(\frac{j}{K}; \theta_1, \theta_2)}{\sum_{j=1}^{K} f(\frac{j}{K}; \theta_1, \theta_2)} \qquad (3)$$

where f is the Beta probability density function.

3.1 Transfer Function Mixed DAta Sampling model (TF-MIDAS)

Expressing MIDAS models in terms of an infinite lag polynomial by introducing a ratio of two finite polynomials is initially suggested in Ghysels et al. (2007).

Our hypothesis is that TF-MIDAS is a more appropiate way to model mixed-frequency data as it is a more general representation than previous DL models, in two ways: (i) removing the truncation error, and (ii) including a MA component in the model. This two new features will be considered in subsection 3.3.

We define the general TF-MIDAS model with N indicators using two equations: (i) the equation that models the relation between y and x:

$$\omega(L)y_t \;=\; \beta_0 \;+\; \sum_{i=1}^{N}\sum_{j=1}^{k} \frac{a_{i,j}(Z)}{b_{i,j}(Z)}\, x_{i,\,t-j} \;+\; \epsilon_t \tag{6a}$$

and, (ii) the equation for the noise:

$$\phi(Z)\epsilon_t \;=\; \theta(Z)a_t, \tag{6b}$$

where $a_j(Z)$ and $b_j(Z)$ are finite lag polynomials, $Z = L^k$; $\phi(Z)$ and $\theta(Z)$ are polynomials of order p and q, respectively, with the usual properties for stationary and invertible ARMA processes.

To estimate TF-MIDAS models, we rearrange the dataset in the following way:

$$\begin{array}{c|cccc|cccc|c|cccc}
y_k^* & x_{1,k-1} & x_{1,k-2} & \dots & x_{1,0} & x_{2,k-1} & x_{2,k-2} & \dots & x_{2,0} & \dots & x_{N,k-1} & x_{N,k-2} & \dots & x_{N,0} \\
y_{2k}^* & x_{1,2k-1} & x_{1,2k-2} & \dots & x_{1,k} & x_{2,2k-1} & x_{2,2k-2} & \dots & x_{2,k} & \dots & x_{N,2k-1} & x_{N,2k-2} & \dots & x_{N,k} \\
\dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\
y_{Tk}^* & x_{1,Tk-1} & x_{1,Tk-2} & \dots & x_{1,Tk-k} & x_{2,Tk-1} & x_{2,Tk-2} & \dots & x_{2,Tk-k} & \dots & x_{N,Tk-1} & x_{N,Tk-2} & \dots & x_{N,Tk-k}
\end{array}$$

where $y_i^* = \omega(L)y_t$ refers to the LF aggregated value of the endogenous variable y; $x_{1,t}$, $x_{2,t}$, ..., $x_{N,t}$ are the HF values of the exogenous variables x_1, x_2, ..., x_N.

Throughout the remaining paper, we will include only one indicator variable x in the model. This is a simplifying assumption and does not affect in any way the results and conclusions derived from the analysis.

For quarterly-annual data (i.e., $k = 3$) and only one indicator the previous equations becomes:

$$\omega(L)y_t \;=\; \beta_0 \;+\; \frac{a_1(Z)}{b_1(Z)}\, x_{t-1} \;+\; \frac{a_2(Z)}{b_2(Z)}\, x_{t-2} \;+\; \frac{a_3(Z)}{b_3(Z)}\, x_{t-3} \;+\; \epsilon_t \tag{7a}$$

$$\phi(Z)\epsilon_t \;=\; \theta(Z)a_t, \tag{7b}$$

where x_{t-1} is a vector formed by the second monthly observation of each quarter, x_{t-2} is formed by the first monthly observation of each quarter, and x_{t-3} is formed by the third monthly observation of each quarter. In this context, the data will be organized in the following way:

$$\begin{array}{c|ccc}
y_3^* & x_{1,2} & x_{1,1} & x_{1,0} \\
y_6^* & x_{1,5} & x_{1,4} & x_{1,3} \\
... & ... & ... & ... \\
y_{3T}^* & x_{1,3T-1} & x_{1,3T-2} & x_{1,3T-3}
\end{array}$$

3.2 Rationale behind TF-MIDAS models

We assume that HF values of variables y and x are generated by a VARMA(p,q) process:

$$\begin{pmatrix} \Phi_{1,1} & \Delta_{1,2} \\ 0 & \Phi_{2,2} \end{pmatrix} \begin{pmatrix} y_t \\ x_t \end{pmatrix} = \begin{pmatrix} \Theta_{1,1} & \Pi_{1,2} \\ 0 & \Theta_{2,2} \end{pmatrix} \begin{pmatrix} e_{y,t} \\ e_{x,t} \end{pmatrix}, \tag{8}$$

where $\Phi_{1,1} = 1 + \phi_{l,1}L + \phi_{l,2}L^2 + ... + \phi_{l,p}L^p$; $\Phi_{2,2} = 1 + \phi_{h,1}L + \phi_{h,2}L^2 + ... + \phi_{h,p}L^p$; $\Delta_{1,2} = \delta_{l,1}L + \delta_{l,2}L^2 + ... + \delta_{l,p}L^p$; $\Theta_{1,1} = 1 + \theta_{l,1}L + \theta_{l,2}L^2 + ... + \theta_{l,q}L^q$; $\Theta_{2,2} = 1 + \theta_{h,1}L + \theta_{h,2}L^2 + ... + \theta_{h,q}L^q$; $\Pi_{1,2} = \pi_{l,1}L + \pi_{l,2}L^2 + ... + \pi_{l,q}L^q$. Throughout the whole paper we assume that $\Phi_{i,i}$ and $\Theta_{i,i}$ polynomials hold the usual conditions for stationarity and invertibility.

We consider several DGPs in order to consider a wide variety of processes: (i) VAR(1); (ii) VMA(1); (iii) VMA(3). In each case we also analyze two aggregation methods for the LF variable: skip-sampling and addition.

As an example, in the next subsection we derive an exact TF-MIDAS representation for a VAR(1) DGP, being y a flow variable (i.e., applying the addition as aggregation method). The same derivations can be found in the Appendix for the other cases.

3.2.1 TF-MIDAS representation of a HF-VAR(1) and a flow variable.

Here we assume monthly values of variables y and x are generated from a VAR(1) process:

$$y_t = \phi_l y_{t-1} + \delta_l x_{t-1} + e_{y,t} \tag{9a}$$
$$x_t = \phi_h x_{t-1} + e_{x,t}. \tag{9b}$$

Shifting (9a), we have:

$$y_{t-k} = \phi_l y_{t-k-1} + \delta_l x_{t-k-1} + e_{y,t-k} \qquad t = 1, 2, 3, ... \tag{10}$$

Considering the previous expression for $k = 1$ and $k = 2$, and substituting them in (9a), yields:

$$y_t = \phi_l^3 y_{t-3} + \phi_l^2 \delta_l x_{t-3} + \phi_l^2 e_{y,t-2} + \phi_l \delta_l x_{t-2} + \phi_l e_{y,t-1} + \delta_l x_{t-1} + e_{y,t} \tag{11}$$

Rearranging terms, we can represent (11) in TF form, expressing the value of y in terms of HF values of variable x and an error term:

$$y_t \;=\; \frac{\delta_l}{1 \;-\; \phi_l^3 L^3}\, x_{t-1} \;+\; \frac{\phi_l \delta_l}{1 \;-\; \phi_l^3 L^3}\, x_{t-2} \;+\; \frac{\phi_l^2 \delta_l}{1 \;-\; \phi_l^3 L^3}\, x_{t-3} \;+\; \frac{1}{1 \;-\; \phi_l L}\, e_{y,t} \quad (12)$$

where $t = 4, 5, 6, ...$ and in order to obtain the last expression we use the factorization $1 - \phi_l^3 L^3 \;=\; (1 - \phi_l L)(1 + \phi_l L + \phi_l^2 L^2)$.

From the expression for y_t in (12) we can deduce equivalent expressions for y_{t-1} and y_{t-2}. Adding these expressions for y_t, y_{t-1} and y_{t-2}, and rearranging terms, we get the aggregated quarterly value of y in the case of a flow variable, denoted by y_t^A:

$$\begin{aligned} y_t^A \;=\; & \frac{\delta_l \;+\; [\phi_l \delta_l (1 + \phi_l)] L^3}{1 \;-\; \phi_l^3 L^3}\, x_{t-1} \;+\; \frac{\delta_l (1 + \phi_l) \;+\; (\phi_l^2 \delta_l) L^3}{1 \;-\; \phi_l^3 L^3}\, x_{t-2} \;+\; \frac{\delta_l (1 + \phi_l + \phi_l^2)}{1 \;-\; \phi_l^3 L^3}\, x_{t-3} \\ & + \frac{1 \;+\; (1 + \phi_l) L \;+\; (1 + \phi_l + \phi_l^2) L^2 \;+\; \phi_l (1 + \phi_l) L^3 \;+\; \phi_l^2 L^4}{1 \;-\; \phi_l^3 L^3}\, e_{y,t} \qquad t = 4, 5, 6, ... \end{aligned} \quad (13)$$

Finally, from (13) we derive the equation that expresses the LF aggregated value of a flow variable y in terms of the HF values of x:

$$\begin{aligned} y_{t_q}^A \;=\; & \frac{\delta_l \;+\; [\phi_l \delta_l (1 + \phi_l)] L}{1 \;-\; \phi_l^3 L}\, x_{2,t_q} \;+\; \frac{\delta_l (1 + \phi_l) \;+\; (\phi_l^2 \delta_l) L}{1 \;-\; \phi_l^3 L}\, x_{1,t_q} \\ & + \frac{\delta_l (1 + \phi_l + \phi_l^2)}{1 \;-\; \phi_l^3 L}\, x_{3,t_q-1} \;+\; \frac{1}{1 \;-\; \phi_l^3 L}\, \eta_{\,t_q} \qquad t_q = 2, 3, 4, ... \end{aligned} \quad (14)$$

where $y_{t_q}^A$ is the aggregated quarterly value of y, which is a flow variable, i.e. $y_{t_q}^A = y_t + y_{t-1} + y_{t-2}$ for $t = 3 * t_q$; x_{2,t_q} is the second monthly value of x for current quarter; x_{1,t_q} is the first monthly value of x for current quarter; x_{3,t_q-1} is the last monthly value of x for the previous quarter; and $\eta_{t_q} \;=\; [1+(1+\phi_l)L+(1+\phi_l+\phi_l^2)L^2+\phi_l(1+\phi_l)L^3+\phi_l^2 L^4]\; e_{y,t}$.

Obviously, η_{t_q} is an autocorrelated noise. We now calculate its principal moments as a function of $e_{y,t}$. Expected value:

$$E[(1 \;+\; (1 + \phi_l) L \;+\; (1 + \phi_l + \phi_l^2) L^2 \;+\; \phi_l (1 + \phi_l) L^3 \;+\; \phi_l^2 L^4)\; e_{y,t}] \;=\; 0, \quad (15)$$

variance:

$$\begin{aligned} & V[(1 \;+\; (1 + \phi_l) L \;+\; (1 + \phi_l + \phi_l^2) L^2 \;+\; \phi_l (1 + \phi_l) L^3 \;+\; \phi_l^2 L^4)\; e_{y,t}] \\ & = \;(3 \;+\; 4\phi_l \;+\; 5\phi_l^2 \;+\; 4\phi_l^3 \;+\; 3\phi_l^4)\; V[e_{y,t}], \end{aligned} \quad (16)$$

and, covariances:

$$\begin{aligned} & Cov[(1 \;+\; (1 + \phi_l) L \;+\; (1 + \phi_l + \phi_l^2) L^2 \;+\; \phi_l (1 + \phi_l) L^3 \;+\; \phi_l^2 L^4)\; e_{y,t}, \\ & \qquad (1 \;+\; (1 + \phi_l) L \;+\; (1 + \phi_l + \phi_l^2) L^2 \;+\; \phi_l (1 + \phi_l) L^3 \;+\; \phi_l^2 L^4)\; e_{y,t-3}] \\ & = \;\phi_l (1 + 2\phi_l + \phi_l^2)\; V[e_{y,t}] \end{aligned} \quad (17a)$$

$$Cov[(1 \; + \; (1+\phi_l)L \; + \; (1+\phi_l+\phi_l^2)L^2 \; + \; \phi_l(1+\phi_l)L^3 \; + \; \phi_l^2 L^4) \; e_{y,t},$$
$$(1 \; + \; (1+\phi_l)L \; + \; (1+\phi_l+\phi_l^2)L^2 \; + \; \phi_l(1+\phi_l)L^3 \; + \; \phi_l^2 L^4) \; e_{y,t-3k}] \tag{17b}$$
$$= \; 0, \quad k = 2,3,4,...$$

From (17a-17b) the autocorrelations are:

$$\rho_1 \; = \; \frac{\phi_l(1+2\phi_l+\phi_l^2)}{3 \; + \; 4\phi_l \; + \; 5\phi_l^2 \; + \; 4\phi_l^3 \; + \; 3\phi_l^4} \tag{18a}$$
$$\rho_k \; = \; 0, \qquad k = 2,3,4,... \tag{18b}$$

Therefore, η_{t_q} presents a MA(1) structure which can be written as:

$$\eta_{t_q} \; = \; (1+\psi L) \; e_{y^A,t_q}, \tag{19}$$

where e_{y^A,t_q} is a white noise process.[4]

Table 1 presents a TF-MIDAS representation for each one of the three DGPs here considered and for both, a stock and a flow variable y.[5]

3.3 Theoretical comparison between TF-MIDAS and U-MIDAS models

Previous transfer function (7a-7b) can also be regarded as a Distributed Lag (DL) model under two restrictions: (i) $\theta(Z) = 1$ and (ii) $b_i(Z) = \phi(Z)$, where lag operator Z has its roots outside the unit circle, in order to ensure stationarity.

Substituting equation (7b) into (7a) and multiplying both sides of the resulting equation by $\phi(Z)$, we get:

$$\phi(Z)\omega(L)y_t \; = \; \beta_0^* \; + \; \omega_1(Z) \; x_{t-1} \; + \; \omega_2(Z) \; x_{t-2} \; + \; \omega_3(Z) \; x_{t-3} \; + \; a_t \tag{20}$$

where $\beta_0^* = \phi(Z)\beta_0$. The expression obtained is equivalent to an U-MIDAS model. This result stands only if the two restrictions mentioned earlier are satisfied. If they are not, then we could derive from equations (7a-7b) this equivalent expression for TF-MIDAS model:

$$\omega(L)y_t \; = \; \beta_0 \; + \; \frac{a_1(Z)}{b_1(Z)} \; x_{t-1} \; + \; \frac{a_2(Z)}{b_2(Z)} \; x_{t-2} \; + \; \frac{a_3(Z)}{b_3(Z)} \; x_{t-3} \; + \; \frac{\theta(Z)}{\phi(Z)} a_t \tag{21}$$

Multiplying both sides of the previous equation by $\phi(Z)$ and re-writing each polynomial quotient as an infinite series, we would arrive at this alternative expression for the TF-MIDAS model:

$$\phi(Z)\omega(L)y_t \; = \; \beta_0^* \; + \; a_1^*(Z) \; x_{t-1} \; + \; a_2^*(Z) \; x_{t-2} \; + \; a_3^*(Z) \; x_{t-3} \; + \; \theta(Z)a_t \tag{22}$$

[4]The value of ψ in (19) can be calculated straightforwardly by using the acf for a MA(1) model: $\rho_1 = \psi/(1+\psi^2)$ and $\rho_k = 0$ for $k = 2,3,\ldots$ Solving the ρ_1 equation for ψ will return two solutions, from where we consider the one satisfying the invertibility condition: $|\psi| < 1$.

[5]Theoretical TF-MIDAS representations, corresponding to a VAR(2) and VARMA(1,1) high-frequency DGPs, are available from the authors upon request.

Table 1: **Theoretical TF-MIDAS models corresponding to selected HF-DGP and different aggregation schemes**

HF-DGP VAR(1)	$y_t = \phi_l y_{t-1} + \delta_l x_{t-1} + e_{y,t}$ $x_t = \phi_h x_{t-1} + e_{x,t}$
TF-MIDAS Stock	$y^A_{t_q} = \frac{\delta_l}{1 - \phi_l^3 L} x_{2,t_q} + \frac{\phi_l \delta_l}{1 - \phi_l^3 L} x_{1,t_q} + \frac{\phi_l^2 \delta_l}{1 - \phi_l^3 L} x_{3,t_q-1} + \zeta_{t_q}$ $(1 - \phi_l^3 L)\, \zeta_{t_q} = e_{y^A,t_q}$
TF-MIDAS Flow	$y^A_{t_q} = \frac{\delta_l + [\phi_l \delta_l (1+\phi_l)] L}{1 - \phi_l^3 L} x_{2,t_q} + \frac{\delta_l (1+\phi_l) + (\phi_l^2 \delta_l) L}{1 - \phi_l^3 L} x_{1,t_q} + \frac{\delta_l (1+\phi_l+\phi_l^2)}{1 - \phi_l^3 L} x_{3,t_q-1} + \zeta_{t_q}$ $(1 - \phi_l^3 L)\, \zeta_{t_q} = (1 + \psi L)\, e_{y^A,t_q}$

HF-DGP VMA(1)	$y_t = e_{y,t} + \theta_l e_{y,t-1} + \delta_l e_{x,t-1}$ $x_t = e_{x,t} + \theta_h e_{x,t-1}$
TF-MIDAS Stock	$y^A_{t_q} = \frac{\delta_l}{1 + \theta_h^3 L} x_{2,t_q} - \frac{\delta_l \theta_h}{1 + \theta_h^3 L} x_{1,t_q} + \frac{\delta_l \theta_h^2}{1 + \theta_h^3 L} x_{3,t_q-1} + \zeta_{t_q}$ $\zeta_{t_q} = e_{y^A,t_q}$
TF-MIDAS Flow	$y^A_{t_q} = \frac{\delta_l + \delta_l \theta_h (\theta_h - 1) L}{1 + \theta_h^3 L} x_{2,t_q} + \frac{\delta_l (1-\theta_h) + \delta_l \theta_h^2 L}{1 + \theta_h^3 L} x_{1,t_q} + \frac{\delta_l (1-\theta_h+\theta_h^2)}{1+\theta_h^3 L} x_{3,t_q-1} + \zeta_{t_q}$ $\zeta_{t_q} = (1 + \psi L)\, e_{y^A,t_q}$

HF-DGP VMA(3)	$y_t = e_{y,t} + \theta_l\, e_{y,t-3} + \delta_l\, e_{x,t-1}$ $x_t = e_{x,t} + \theta_h\, e_{x,t-1}$
TF-MIDAS Stock	$y^A_{t_q} = \frac{\delta_l}{1 + \theta_h^3 L} x_{2,t_q} - \frac{\delta_l \theta_h}{1 + \theta_h^3 L} x_{1,t_q} + \frac{\delta_l \theta_h^2}{1 + \theta_h^3 L} x_{3,t_q-1} + \zeta_{t_q}$ $\zeta_{t_q} = (1 + \theta_l L)\, e_{y^A,t_q}$
TF-MIDAS Flow	$y^A_{t_q} = \frac{\delta_l + \delta_l \theta_h (\theta_h - 1) L}{1 + \theta_h^3 L} x_{2,t_q} + \frac{\delta_l (1-\theta_h) + \delta_l \theta_h^2 L}{1 + \theta_h^3 L} x_{3,t_q} + \frac{\delta_l (1-\theta_h+\theta_h^2)}{1 + \theta_h^3 L} x_{1,t_q-1} + \zeta_{t_q}$ $\zeta_{t_q} = (1 + \theta_l L)\, e_{y^A,t_q}$

where $\beta_0^* = \phi(Z)\beta_0$; $a_i^*(Z) = \sum_{j=0}^{\infty} a *_{i,j} L^j$, $i = 1, 2, 3, \ldots$.

Therefore using U-MIDAS model when the two aforementioned conditions are not satisfied would imply truncating the lag polynomials in the previous expression, and therefore working with an approximation of the true model. How good this approximation results depends on the specific characteristics of the unobservable HF-DGP. For example, consider the following TF-MIDAS model:

$$\omega(L)y_t = \beta_0 + \frac{a_{0,1} + a_{1,1}L}{1 - b_1 L} x_{t-1} + \frac{a_{0,2} + a_{1,2}L}{1 - b_2 L} x_{t-2} + \frac{a_{0,3} + a_{1,3}L}{1 - b_3 L} x_{t-3} + \epsilon_t \quad (23)$$

Assuming $0 < b_j < 1$, the quotient $(a_{0,j} + a_{1,j}L)/(1 - b_j L)$ can be expanded to $a_{0,j} + (a_{1,j} + a_{0,j}b_j)L + \delta_j(a_{1,j} + a_{0,j}b_j)L^2 + b_j^2(a_{1,j} + a_{0,j}b_j)L^3 + \ldots$, which is a polynomial in L of infinite order. So, the model in TF representation can be rewritten as:

$$\omega(L)y_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \ldots + \epsilon_t, \qquad t = 3, 6, \ldots, 3T, \quad (24)$$

where $\beta_n = a_{0,n}$, for $n = 1, 2, 3$; $\beta_{n+3} = (a_{1,n} + a_{0,n}b_n)$, for $n = 1, 2, 3$; $\beta_{n+3i+3} = b_n^i(a_{1,n} + a_{0,n}b_n)$, for $n = 1, 2, 3, \ldots$ and $i = 1, 2, 3, \ldots$

If instead of a TF representation, we opt for a DL one, it will be specified in terms of a finite-order polynomial, e.g., a third-order polynomial:

$$\omega(L)y_t = \beta_0 + \beta_1 x_{t-1} + \beta_2 x_{t-2} + \beta_3 x_{t-3} + \epsilon_t. \quad (25)$$

Besides the difference in the specification of the error term structure, a crucial difference between both models is the truncation of the lag polynomial in the DL representation. This truncation implies discarding the terms of higher order in the lag polynomial. In our example it implies to leave out the terms of forth and higher order: $\beta_4 x_{t-4} + \beta_5 x_{t-5} + \ldots$

For some HF-GDPs, the use of a DL approximation will result then in different nowcasts and forecasts with respect to the ones obtained with a TF representation.

On the other hand, although the TF-MIDAS model is more accurate and parsimonious, it has to be estimated by non-linear estimation method, such as Maximum-Likelihood (ML). This, of course, represents a bigger computational challenge than fitting UMIDAS by least squares (LS) procedures. Fortunately, advances in computing power have significantly reduced this potential drawback.

4 Nowcast performance evaluation

In this section we carry out several Monte Carlo simulation exercises in order to test the nowcasting performance of TF-MIDAS models against U-MIDAS models for different HF-DGP.

4.1 Simulation design

The simulation design applied in this paper closely follows that one considered in Foroni et al. (2015), making the results comparable. We simulate three DGPs: (i) HF-VAR(1), (ii) HF-VMA(1) and (ii) HF-VMA(3). We consider a series of parameter combinations that seek to represent a wide range of DGPs, with different degrees of persistence and correlation between high-frequency and low-frequency variables. Including such a variety of DGPs aims to provide more robustness to the conclusions.

Following Foroni et al. (2015), y_t and x_t are initially simulated for all $t = 1, ..., (T + ES) \times k$, where t is the HF time index and k denotes the sampling frequency of the LF variable (e.g., for a quarterly observed variable y and a monthly indicator x, $k = 3$).

The number of observations simulated for both variables are used to estimate the nowcasting models is $T \times k = 300$. In order to perform a nowcast comparison we also consider an Evaluation Sample (ES), so both variables are also simulated $ES \times k$ periods ahead. The size of the ES is set equal to T/2=50.

Once all the values for both variables are simulated, the ones corresponding to the LF variable are aggregated. We consider two different aggregation rules, depending if the variable is a stock or a flow. In the former case, we aggregate the values applying the so-called *skip-sampling* procedure, which consists in considering the aggregated values as the values of the LF variable y corresponding only to $t = k, 2k, ..., (T + ES) \times k$; i.e., $\omega(L)y_t = y_t$ for $t = k, 2k, ..., (T + ES) \times k$. In the case of a flow variable, we obtain the quarterly aggregated value adding the corresponding monthly values; i.e., $\omega(L)y_t = y_t + y_{t-1} + y_{t-2}$.

In the following lines we present the main characteristics of each DGP considered.

DGP I: HF-VAR(1)

Firstly, we consider a HF bivariate VAR process of order 1, defined by Equation (8) with $\Phi_{1,1} = 1 + \phi_{l,1}L$, $\Phi_{2,2} = 1 + \phi_{h,1}L$, $\Delta_{1,2} = \delta_{l,1}$, $\Pi_{1,2} = 0$ and $\Theta_{1,1} = \Theta_{2,2} = 1$. Error terms $e_{y,t}$ and $e_{x,t}$ are sampled independently from the normal distribution with mean 0 and variance chosen such that the unconditional variance of y is equal to 1, given the specifications of the other parameters.

Parameters $\phi_{l,1}$ and $\phi_{h,1}$, which determine the persistance of both series, are chosen to represent three types of processes with different levels of persistence: low ($\phi_{l,1} = \phi_{h,1} = 0.1$), medium ($\phi_{l,1} = \phi_{h,1} = 0.5$) and high ($\phi_{l,1} = \phi_{h,1} = 0.9$).

Parameter $\delta_{l,1}$ reflects the dependence of variable y on the past value of variable x. Its value is chosen such that stationarity of the series is ensured. Following Foroni et al. (2015), which in turn follows Ghysels and Valkanov (2006), values considered for $\delta_{l,1}$ are 0.1, 0.5 and 1.0. As it is usually assumed in the literature, high-frequency variable affects low-frequency variable, but not the other way round.

DGP II: HF-VMA(1)

The first HF bivariate VMA process is of order 1, inferior to the sampling frequency ($k = 3$), so that in the case of a stock variable, previous quarter LF variable's MA component does not affect its current quarter value. It is defined by equation (8), with $\Phi_{1,1} = \Phi_{2,2} = 1$, $\Delta_{1,2} = 0$, $\Pi_{1,2} = \pi_{l,1}L$, $\Theta_{1,1} = 1 - \theta_{l,1}L$, and $\Theta_{2,2} = 1 - \theta_{h,1}L$.

We consider different combinations of the parameters values, in order to include a wide variety of DGPs. For θ_l and θ_h we consider values $\{0.5, 0.7, 0.9\}$, for δ_l the values $\{0.1, 0.5, 1\}$ and we set $\delta_h = 0$. We distinguish GDPs with $\theta_l = 0$, i.e., with no MA component, with $\theta_h = 0$, and with both parameters different from 0.

DGP III: HF-VMA(3)

The second HF bivariate VMA process is of order 3, equal to the sampling frequency, so that MA component does have an impact on LF variable's current value. This process is also defined by equation (8) with $\Phi_{1,1} = \Phi_{2,2} = 1$, $\Delta_{1,2} = 0$, $\Pi_{1,2} = \pi_{l,1}L$ and $\Theta_{2,2} = 1 - \theta_{h,1}L$. However, now $\Theta_{1,1} = 1 - \theta_{l,1}L^3$.

Once again we consider different combinations of the parameters values in order to provide more robustness to the analysis. We consider for θ_l and θ_h the values $\{0, -0.5, -0.7, -0, 9\}$, for δ_l the values $\{0.1, 0.5, 1\}$ and θ_h is set to 0.

4.2 Models under comparison

Foroni et al. (2015) show that U-MIDAS beats the standard MIDAS model when the difference in sampling frequency is not high, as it is in our case. We therefore consider U-MIDAS (and not standard MIDAS) as the reference to evaluate the nowcasting performance of TF-MIDAS model. We then estimate and compare the nowcasting performance of two model classes.

(i) U-MIDAS, defined by the following equation:

$$\omega(L)y_t \;=\; \mu_0 + c_1\omega(L)y_{t-k} + \delta(L)x_{t-1} + \epsilon_t, \quad t = k, 2k, ..., k(T_q - 1); \tag{26}$$

where $\delta(L) = \sum_{j=0}^{K} \delta_j L^j = \delta_0 + \delta_1 L + ... + \delta_K L^K$ and in our case $k = 3$.

We do not impose the common factor restriction as in Clements and Galvão (2008). The coefficients μ_0, c_1, δ_0, ..., δ_K are estimated by LS. We use BIC to determine the lag order K. Following Foroni et al. (2015), we consider two different cases, including (i) $K_{max} = k$, i.e., up to 3 lags of x_t, and (ii) $K_{max} = 4k$, i.e., up to 12 lags of x_t.

Once we have estimated the parameters, U-MIDAS nowcast is computed as,

$$\omega(L)\hat{y}_{T\times k+es\times k\ |T\times k+es\times k-1} \;=\; \hat{\mu}_0 + \hat{c}_1\omega(L)y_{T\times k+es\times k-k} + \hat{\delta}(L)\ x_{T\times k+es\times k-1}, \tag{27}$$

(ii) TF-MIDAS, defined by the equation:

$$\omega(L)y_t \;=\; \beta_0 + \frac{a_1(Z)}{b_1(Z)}\ x_{t-1} + \frac{a_2(Z)}{b_2(Z)}\ x_{t-2} + \frac{a_3(Z)}{b_3(Z)}\ x_{t-3} + \frac{\theta(Z)}{\phi(Z)}\ a_t, \quad t = k, 2k, ..., kT_q \tag{28}$$

Several variants of this model are considered, including different orders of the lag polynomials. In each estimation stage, as in the UMIDAS case, the model chosen is the one with the least BIC, which it is then used to compute the nowcast.

The parameters are estimated by exact ML through the Kalman Filter (see, Casals et al., 2016, Chapters 5 and 6). Then, the corresponding nowcast is obtained from the following calculation:

$$\begin{aligned}\omega(L)\hat{y}_{kT_q+es \mid kT_q+es-1} &= \hat{\beta}_0 + \frac{\hat{a}_1(Z)}{\hat{b}_1(Z)}\, x_{kT_q+es-1} + \frac{\hat{a}_2(Z)}{\hat{b}_2(Z)}\, x_{kT_q+es-2} + \\ &\quad \frac{\hat{a}_3(Z)}{\hat{b}_3(Z)}\, x_{kT_q+es-3} + \frac{\hat{\theta}(Z)}{\hat{\phi}(Z)}\, \hat{a}_{kT_q+es} \qquad es = k, 2k, ..., ES \end{aligned} \tag{29}$$

where $\hat{\theta}(Z) = \theta_1 Z + \theta_2 Z^2 + ... + \theta_q Z^q$.[6]

4.3 Nowcast performance evaluation

4.3.1 Evaluation procedure

The aggregated simulated values for the LF variable corresponding to the ES are used as the actual values to be compared with the nowcasted values. We assume that the information set available for nowcasting consists in values up to period (T + es - 1) $\times$ k, with $es = 1, ..., ES$, for the LF variable, and up to period (T + es - 1) $\times$ k + k - 1 for the HF variable. These values are considered known in each nowcasting exercise and are used to estimate TF-MIDAS and U-MIDAS models.

Forecasts of y_t are then computed one LF period ahead (or equivalently k HF periods ahead) for each date in the evaluation sample, conditional on HF information available within the LF forecast period: $\hat{y}_{T\times k+es\times k \mid T\times k+es\times k-1}$. The corresponding nowcast error is calculated as $\hat{y}_{T\times k+es\times k \mid T\times k+es\times k-1}$ - $y_{T\times k+es\times k}$.

In line with the literature, the indicator used to compare the out-of-sample nowcasting performance of the alternative mixed-frequency models is the Mean Square Forecasting Error (MSFE) over the evaluation sample. That is, for each replication r we have:

$$MSFE_r = \frac{1}{ES} \sum_{es=1}^{ES} (\hat{y}_{T\times k+es\times k \mid T\times k+es\times k-1} - y_{T\times k+es\times k})^2, \tag{30}$$

where $r = 1, ..., R$. In this research the total number of replications (R) amounts to 500.

Therefore, as in Foroni et al. (2015), for each replication we need to estimate the model and then obtain the one HF period ahead nowcast 50 times: one for each quarter of the evaluation sample. In each replication the alternative mixed-frequency models are compared through the ratio of their MSFE: $MSFE_{\ TF-MIDAS}/MSFE_{\ U-MIDAS}$.

In order to perform the calculations we use the MatLab toolboxes E4 (Casals et al., 2016) and Midas (Ghysels, E. and collaborators, 2017). Main descriptive statistics of this ratio are presented in the following tables for each one of the GDPs considered and the cases of stock and flow variables.

[6]Notice that the polynomial $\hat{\theta}(Z)$ does not include the unit term as $\hat{a}_{kT_q+es}$ is not known at period $kT_q + es$.

4.3.2 Monte Carlo nowcast comparison results

The relative nowcast performance of TF-MIDAS respect to UMIDAS is presented in Tables 2-4.

In the case of a HF-VAR(1) DGP, we do not find any advantage in using TF-MIDAS, whether LF variable y is a stock or a flow, considering different parameter combinations and both, $K_{max} = 3$ and $K_{max} = 12$ (see, Tables 2A-2B and Figure 1). This result is expected in the case of a stock variable as TF-MIDAS model expression is observationally equivalent to the corresponding UMIDAS, including one period-lag of the dependent variable. In the case of a flow variable, one could expect TF-MIDAS would provide some nowcasting performance improvement, as it accommodates an existing MA component to the error term. However, it appears that this effect is so low in this case that no improvement is observed. The MA parameter ψ in TF-MIDAS model is a direct function of the value of the AR parameter ϕ_l in the DGP. Even considering a highly persistent LF process, e.g. $\phi_l = 0.9$, parameter ψ reaches a small value of approximately 0.22. So, most of times, the BIC criterion chooses a TF-MIDAS model that does not include the MA term and the potentially benefit of using this model disappears.[7]

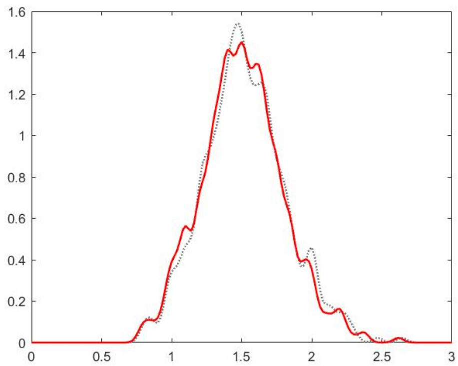
In the case of a HF-VMA(1) DGP, we do not find that TF-MIDAS provides a significant improvement in nowcasting performance relative to UMIDAS for a stock variable (Table 3A).

However, when a flow aggregation scheme is considered, TF-MIDAS shows a significant improvement of its relative nowcasting performance that goes, in median terms, up to 56% when $K_{max} = 3$ and 22%, when $K_{max} = 12$ (see, Table 3B and Figure 1). The explanation of this results mainly comes from the MA structure of the DGP, as we are considering a HF-VMA DGP of order 1, less than the sampling frequency k, which is 3. This causes that in the stock-variable case the previous quarter LF error term ($e_{y,t-3}$) has no effect on the current value of the LF variable (i.e., the mixed-frequency model has no MA component). But in the flow-variable case, when addition is applied as aggregation method and $\theta_l \neq 0$, this effect endures and a non-negligible MA component appears in the mixed-frequency model.

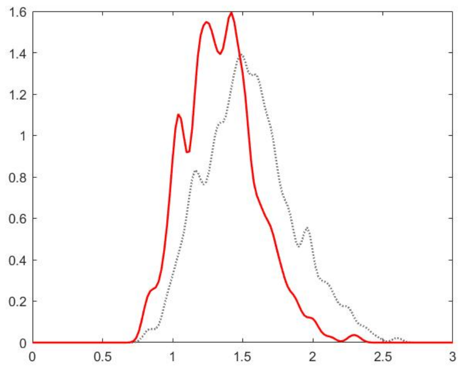
The greater difference in nowcasting performance in favor of TF-MIDAS is observed when TF-MIDAS' two potential sources of relative advantage occur: (i) a polynomial quotient expression, which damages UMIDAS particularly when K_{max} is short, and, (ii) a MA component in the mixed-frequency model appears. In the DGP context, this happens for $\theta_l \neq 0$ and $\theta_h \neq 0$. Besides, results show that TF-MIDAS relative performance improves as the value of δ_l and/or θ_l or θ_h increase (see, Figure 1b). Last, as said before, TF-MIDAS advantage tends to reduce when a greater value K_{max} is chosen (e.g., $K_{max} = 12$). This is a reasonable result, as considering more lags of the indicator makes UMIDAS model to better approximate TF-MIDAS' coefficient quotients.

Finally, as predicted by theory developed in Section 3.2, in the case of a HF-VMA(3)

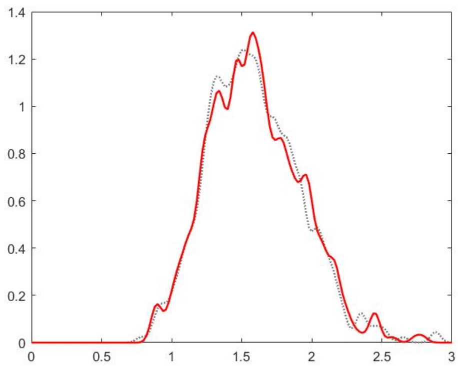
[7] However, some alternative simulations have been performed using the AIC criterion to select the best forecasting model. As this criterion overweight the goodness-of-fit against the parsimony, relative to BIC, the criterion does choose the TF-MIDAS with an MA term most of the times, leading to the expected nowcasting performance improvement with respect to UMIDAS models. The results of these simulations are available from the authors upon request.



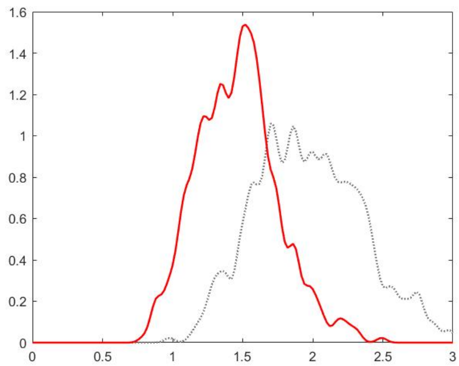
(a) $\theta_l = 0.7$; $\theta_h = 0$; $\delta_l = 0.5$
Aggregation: skip-sampling



(b) $\theta_l = 0.7$; $\theta_h = 0$; $\delta_l = 0.5$
Aggregation: addition



(c) $\theta_l = 0.7$; $\theta_h = 0.7$; $\delta_l = 0.5$
Aggregation: skip-sampling



(d) $\theta_l = 0.7$; $\theta_h = 0.7$; $\delta_l = 0.5$
Aggregation: addition

Figure 1: Kernel density estimates for $MSFE_{TF-MIDAS}$ (red) and $MSFE_{UMIDAS}$ (grey). DGP: HF-VMA(1). $K_{max} = 3$. Rows: same parameter values. Columns: same aggregation scheme.

DGP, the TF-MIDAS advantage is observed whether a stock or a flow variable y is considered (Tables 4A-4B and Figure 3). The benefits in terms of relative performance go up to 23% for $K_{max} = 3$ and 30% for $K_{max} = 12$ when the aggregation method is skip-sampling (stock), and to 36% and 19% for $K_{max} = 3$ and $K_{max} = 12$, respectively, when the aggregation method is addition (flow). Notice that for a stock aggregation scheme, the only case where no improvement is reported is when parameter θ_l is zero.

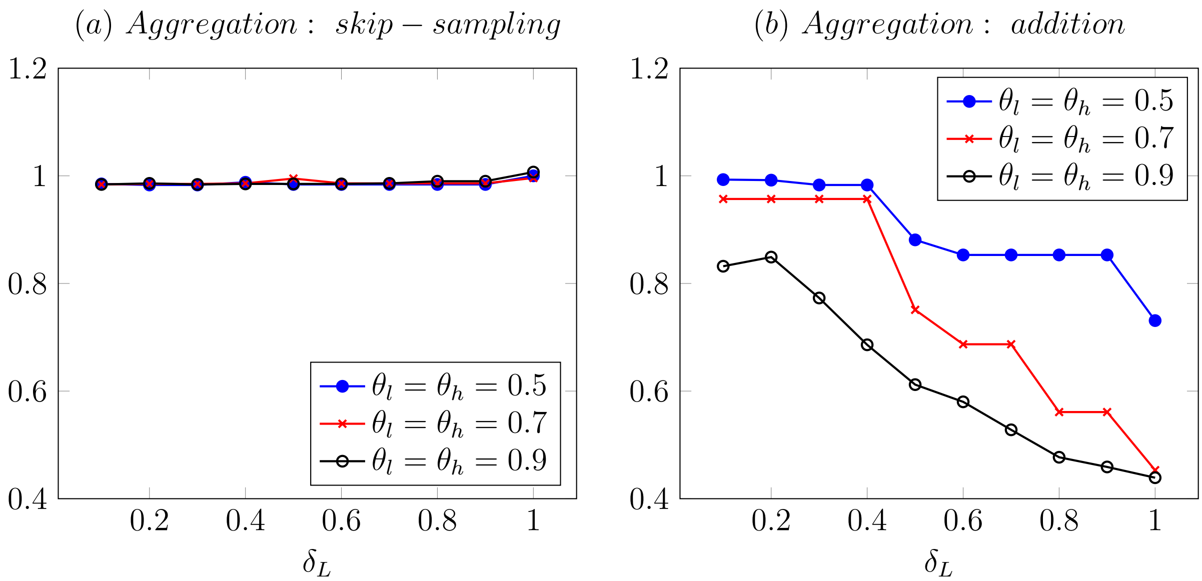
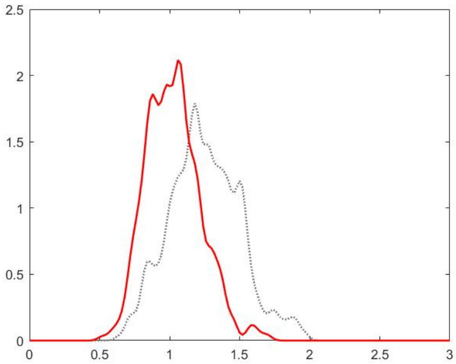
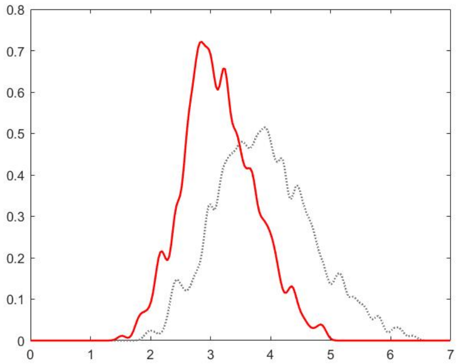


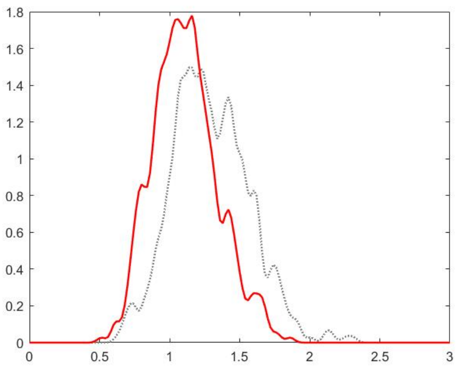
Figure 2: $Median\ (\frac{MSFE\ _{TF-MIDAS}}{MSFE\ _{UMIDAS}})$ - DGP: HF-VMA(1) - $K_{max} = 3$



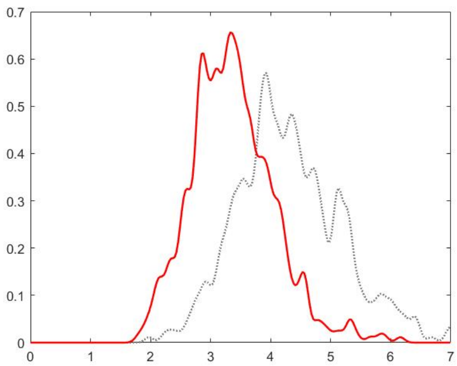
(a) $\theta_l = 0.9$; $\theta_h = 0$; $\delta_l = 0.5$
Aggregation: skip-sampling



(b) $\theta_l = 0.9$; $\theta_h = 0$; $\delta_l = 0.5$
Aggregation: addition



(c) $\theta_l = 0.9$; $\theta_h = 0.9$; $\delta_l = 0.5$
Aggregation: skip-sampling



(d) $\theta_l = 0.9$; $\theta_h = 0.9$; $\delta_l = 0.5$
Aggregation: addition

Figure 3: Kernel density estimates for $MSFE_{TF-MIDAS}$ (red) and $MSFE_{UMIDAS}$ (grey). DGP: HF-VMA(3). $K_{max} = 3$. Rows: same parameter values. Columns: same aggregation scheme.

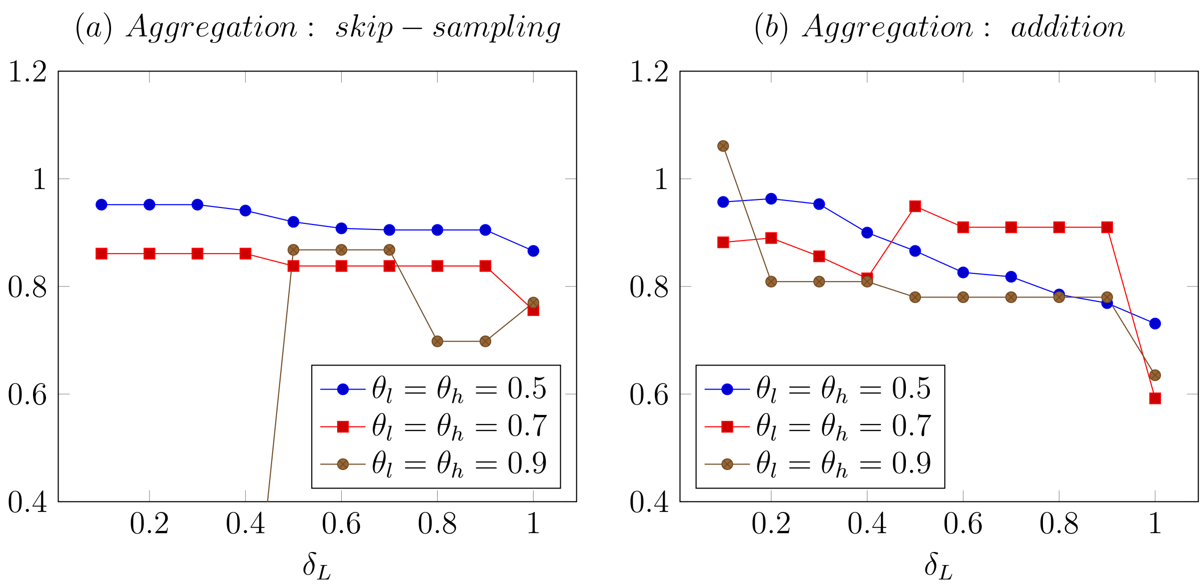


Figure 4: $Median\ (\frac{MSFE_{\ TF-MIDAS}}{MSFE_{\ UMIDAS}})$ - DGP: HF-VMA(3) - $K_{max} = 3$

5 Conclusions

This paper deals with the mixed-frequency modeling problem from a new perspective. Instead of drawing upon the common distributed lag polynomial model, we use a transfer function representation to develop a new type of models, named TF-MIDAS.

We describe the model and derive the theoretical TF-MIDAS implied by the high-frequency (bivariate) VARMA family models, depending on the aggregation scheme. This exact correspondence leads to two potential gains in terms of nowcasting and forecasting performance against, for instance, the common alternative UMIDAS. This is so because TF-MIDAS adds to current UMIDAS new terms to capture: (i) different infinite low-frequency variable responses to shocks in the high-frequency variable, and (ii) a MA structure in the error term.

As predicted by the theoretical development, an extensive Monte Carlo simulation exercise confirms that TF-MIDAS beats currently available UMIDAS models in terms of out-of-sample nowcasting performance for HF-VMA(1) and HF-VMA(3) processes, while their performance remains very similar for HF-VAR(1). These results are robust to different parameter specifications.

An advantage of TF-MIDAS models is that they lie on common transfer functions solid specification and estimation methods (see, e.g., Box and Jenkins, 1976). Instead, the estimation is computationally more expensive than alternatives based on least squares methods.

Table 2: Out-of-sample MSFE of TF-MIDAS relative to U-MIDAS (DGP: HF-VAR(1)

A) Aggregation method: skip-sampling

Parameters			Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
ρ	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0.1	0.1	0	0.996	0.962	0.982	0.993	1.009	1.030	0.995	0.963	0.982	0.994	1.009	1.029
0.1	0.5	0	0.993	0.958	0.978	0.991	1.008	1.033	0.992	0.955	0.974	0.991	1.009	1.032
0.1	1.0	0	0.996	0.957	0.975	0.995	1.014	1.035	0.995	0.956	0.974	0.993	1.013	1.034
0.5	0.1	0	1.000	0.964	0.981	0.996	1.015	1.041	0.999	0.959	0.980	0.996	1.015	1.039
0.5	0.5	0	1.010	0.952	0.981	1.008	1.038	1.071	1.008	0.949	0.979	1.006	1.036	1.067
0.5	1.0	0	1.023	0.947	0.985	1.020	1.058	1.099	1.019	0.943	0.977	1.017	1.056	1.096
0.9	0.1	0	1.010	0.948	0.976	1.004	1.041	1.085	1.009	0.947	0.976	1.004	1.038	1.084
0.9	0.5	0	1.040	0.961	0.992	1.033	1.075	1.137	1.037	0.958	0.991	1.030	1.074	1.131
0.9	1.0	0	1.285	1.003	1.062	1.187	1.369	1.650	1.281	0.998	1.060	1.179	1.369	1.646

B) Aggregation method: addition

Parameters			Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
ρ	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0.1	0.1	0	1.022	0.956	0.986	1.010	1.050	1.100	1.021	0.957	0.983	1.009	1.049	1.098
0.1	0.5	0	1.075	0.968	1.008	**1.066**	1.125	1.198	1.070	0.966	1.008	**1.062**	1.122	1.191
0.1	1.0	0	1.036	0.950	0.991	1.030	1.073	1.124	1.031	0.949	0.988	1.027	1.067	1.121
0.5	0.1	0	1.017	0.951	0.981	1.015	1.049	1.091	1.016	0.949	0.978	1.013	1.047	1.092
0.5	0.5	0	1.051	0.949	0.991	1.042	1.100	1.171	1.045	0.944	0.987	1.034	1.094	1.159
0.5	1.0	0	0.989	0.872	0.920	0.981	1.046	1.112	1.004	0.916	0.951	0.996	1.054	1.102
0.9	0.1	0	0.983	0.879	0.925	0.979	1.040	1.099	0.981	0.877	0.922	0.978	1.035	1.095
0.9	0.5	0	1.004	0.828	0.903	0.992	1.099	1.192	1.041	0.898	0.963	1.031	1.114	1.197
0.9	1.0	0	1.299	0.745	0.870	1.135	1.598	2.294	1.659	0.970	1.090	1.433	1.963	2.997

In bold rows where median MSFE of TF-MIDAS is at least 5% better than U-MIDAS.

Table 3: Out-of-sample MSFE of TF-MIDAS relative to U-MIDAS (DGP: HF-VMA(1)

A) Aggregation method: skip-sampling

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0	-0.5	0.1	0	0.987	0.947	0.968	0.986	1.007	1.030	0.983	0.945	0.965	0.983	1.004	1.024
0	-0.7	0.5	0	1.012	0.962	0.981	1.003	1.037	1.078	0.986	0.936	0.961	0.984	1.009	1.033
0	-0.7	1	0	1.036	0.940	0.980	1.015	1.071	1.154	1.022	0.913	0.964	1.000	1.066	1.147
-0.5	0	0.1	0	0.991	0.956	0.971	0.986	1.010	1.039	0.988	0.949	0.969	0.985	1.007	1.031
-0.7	0	0.5	0	0.990	0.942	0.969	0.988	1.005	1.038	0.986	0.939	0.964	0.986	1.002	1.028
-0.9	0	1	0	0.988	0.945	0.967	0.985	1.005	1.030	0.983	0.943	0.964	0.983	1.002	1.023
-0.5	-0.5	0.1	0	0.988	0.947	0.968	0.985	1.006	1.037	0.982	0.940	0.964	0.981	1.001	1.027
-0.7	-0.7	0.5	0	1.004	0.954	0.974	0.995	1.029	1.071	0.982	0.936	0.960	0.978	1.003	1.032
-0.9	-0.9	1	0	1.016	0.937	0.975	1.007	1.055	1.096	0.993	0.901	0.946	0.985	1.035	1.083

B) Aggregation method: addition

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0	-0.5	0.1	0	0.996	0.953	0.973	0.989	1.011	1.048	0.988	0.943	0.967	0.986	1.005	1.034
0	-0.7	0.5	0	1.004	0.936	0.972	0.999	1.038	1.103	1.029	0.931	0.971	1.011	1.088	1.160
0	-0.7	1	0	0.798	0.635	0.695	**0.781**	0.904	1.013	1.010	0.880	0.930	0.983	1.054	1.181
-0.5	0	0.1	0	0.941	0.851	0.886	**0.941**	0.990	1.031	0.940	0.852	0.891	**0.941**	0.987	1.024
-0.7	0	0.5	0	0.880	0.779	0.820	**0.869**	0.934	0.990	0.879	0.780	0.821	**0.870**	0.931	0.988
-0.9	0	1	0	0.735	0.586	0.631	**0.708**	0.825	0.936	0.829	0.666	0.718	**0.796**	0.910	1.050
-0.5	-0.5	0.1	0	0.997	0.940	0.963	0.993	1.020	1.068	0.989	0.933	0.957	0.988	1.012	1.057
-0.7	-0.7	0.5	0	0.752	0.602	0.676	**0.751**	0.827	0.897	0.910	0,785	0.840	**0.904**	0.969	1.044
-0.9	-0.9	1	0	0.480	0.291	0.359	**0.439**	0.586	0.733	0.879	0.573	0.644	**0.780**	1.093	1.329

In bold rows where median MSFE of TF-MIDAS is at least 5% better than U-MIDAS.

Table 4: Out-of-sample MSFE of TF-MIDAS relative to U-MIDAS (DGP: HF-VMA(3)

A) Aggregation method: skip-sampling

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0	-0.5	0.1	0	0.998	0.970	0.985	0.994	1.006	1.025	0.997	0.969	0.983	0.993	1.006	1.025
0	-0.7	0.5	0	1.000	0.960	0.980	0.997	1.015	1.041	0.960	0.642	0760	**0.926**	1.123	1.324
0	-0.7	1	0	1.010	0.939	0.975	1.004	1.048	1.093	1.069	0.702	0.834	1.020	1.237	1.532
0	-0.9	1	0	0.941	0.814	0.867	**0.941**	1.005	1.060	0.940	0.819	0.872	**0.930**	1.011	1.069
-0.5	0	0.1	0	0.954	0.874	0.909	**0.950**	0.991	1.038	0.953	0.871	0.908	**0.950**	0.992	1.034
-0.7	0	0.1	0	0.950	0.864	0.905	**0.949**	0.993	1.037	0.950	0.861	0.904	**0.950**	0.992	1.038
-0.7	0	0.5	0	0.776	0.656	0.710	**0.768**	0.830	0.899	0.806	0.696	0.740	**0.796**	0.863	0.933
-0.9	0	0.5	0	0.828	0.725	0.769	**0.816**	0.889	0.965	0.827	0.725	0.767	**0.818**	0.885	0.954
-0.9	0	1	0	0.699	0.577	0.624	**0.690**	0.763	0.843	0.889	0.552	0.685	**0.843**	1.043	1.247
-0.5	-0.5	0.1	0	0.956	0.876	0.910	0.952	0.998	1.044	0.955	0..874	0.912	0.952	0.999	1.043
-0.7	-0.7	0.1	0	0.873	0.767	0.812	**0.861**	0.926	1.000	0.871	0.766	0.808	**0.859**	0.926	1.000
-0.7	-0.7	0.5	0	0.849	0.731	0.780	**0.838**	0.912	0.973	0.851	0.735	0.787	**0.840**	0.908	0.976
-0.9	-0.9	0.5	0	0.870	0.727	0.798	**0.868**	0.938	1.005	0.873	0.739	0.805	**0.872**	0.937	0.989
-0.9	-0.9	1	0	0.782	0.658	0.710	**0.770**	0.851	0.912	0.733	0.488	0.583	**0.707**	0.856	1.011

B) Aggregation method: addition

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0	-0.5	0.1	0	0.997	0.946	0.972	0.994	1.020	1.056	0.996	0.942	0.971	0.994	1.020	1.056
0	-0.7	0.5	0	0.941	0.820	0.879	**0.943**	0.999	1.058	0.972	0.896	0.929	0.973	1.007	1.051
0	-0.7	1	0	0.996	0.693	0.772	**0.892**	1.183	1.473	1.267	0.926	0.985	1.066	1.487	1.878
0	-0.9	1	0	0.749	0.565	0.645	**0.732**	0.838	0.935	0.994	0.787	0.853	0.962	1.088	1.235
-0.5	0	0.1	0	1.219	1.039	1.104	**1.177**	1.298	1.432	1.216	1.033	1.102	**1.178**	1.293	1.426
-0.7	0	0.1	0	0.953	0.870	0.907	**0.946**	0.998	1.041	0.952	0.870	0.905	**0.946**	0.997	1.038
-0.7	0	0.5	0	0.820	0.712	0.754	**0.810**	0.873	0.955	0.819	0.701	0.744	**0.815**	0.876	0.954
-0.9	0	0.5	0	0.817	0.697	0.745	**0.809**	0.879	0.945	0.818	0.702	0.751	**0.811**	0.874	0.938
-0.9	0	1	0	0.854	0.611	0.682	**0.810**	0.992	1.166	0.892	0.645	0.724	**0.849**	1.023	1.198
-0.5	-0.5	0.1	0	0.961	0.861	0.915	0.957	1.004	1.057	0.959	0.860	0.912	0.955	1.003	1.053
-0.7	-0.7	0.1	0	0.888	0.777	0.828	**0.882**	0.943	1.013	0.887	0.775	0.826	**0.882**	0.942	1.009
-0.7	-0.7	0.5	0	0.956	0.877	0.908	**0.949**	0.997	1.041	0.954	0.873	0.908	**0.946**	0.995	1.039
-0.9	-0.9	0.5	0	0.786	0.648	0.706	**0.780**	0.860	0.934	0.828	0.691	0.749	**0.823**	0.901	0.979
-0.9	-0.9	1	0	0.636	0.429	0.519	**0.635**	0.746	0.851	0.838	0.592	0.668	**0.829**	0.963	1.101

In bold rows where median MSFE of TF-MIDAS is at least 5% better than U-MIDAS.

References

Bai, J., Ghysels, E., and Wright, J. H. (2013). State space models and MIDAS regressions. *Econometric Reviews*, 32(7):779–813.

Box, G. E. P. and Jenkins, G. M. (1976). *Time series analysis: forecasting and control.* Holden-Day.

Casals, J., García-Hiernaux, A., Jerez, M., and Sotoca, S. (2016). *State-space methods for time series analysis: theory, applications and software.* Chapman & Hall.

Castle, J. and Hendry, D. (2013). Forecasting and nowcasting macroeconomic variables: A methodological overview. Economics Series Working Paper 674, University of Oxford, Department of Economics.

Clements, M. P. and Galvão, A. B. (2008). Macroeconomic forecasting with mixed-frequency data: Forecasting output growth in the United States. *Journal of Business & Economic Statistics*, 26(4):546–554.

Clements, M. P. and Galvão, A. B. (2009). Forecasting US output growth using leading indicators: an appraisal using MIDAS models. *Journal of Applied Econometrics*, 24(7):1187–1206.

Duarte, C., Rodrigues, P. M. M., and Rua, A. (2017). A mixed frequency approach to the forecasting of private consumption with ATM/POS data. *International Journal of Forecasting*, 33(1):61–75.

Foroni, C., Marcellino, M., and Schumacher, C. (2012). U-MIDAS: MIDAS regressions with unrestricted lag polynomials. CEPR Discussion Paper 8828.

Foroni, C., Marcellino, M., and Schumacher, C. (2015). Unrestricted mixed data sampling (MIDAS): MIDAS regressions with unrestricted lag polynomials. *Journal of the Royal Statistical Society Series A*, 178(1):57–82.

Foroni, C. and Marcellino, M. G. (2013). A survey of econometric methods for mixed-frequency data. Norges Bank Working Paper.

Ghysels, E. (2014). Matlab toolbox for mixed sampling frequency data analysis using MIDAS regression models. Technical report.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2002). The MIDAS touch: Mixed data sampling regression models. Working paper, UNC and UCLA.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2003). Predicting volatility: Getting the most out of return data sampled at different frequencies. Anderson School of Management working paper and UNC Department of Economics working paper.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics*, 76(3):509 – 548.

Ghysels, E., Santa-Clara, P., and Valkanov, R. (2006). Predicting volatility: getting the most out of return data sampled at different frequencies. *Journal of Econometrics*, 131(1):59–95.

Ghysels, E., Sinko, A., and Valkanov, R. (2007). MIDAS regressions: Further results and new directions. *Econometric Reviews*, 26(1):53–90.

Ghysels, E. and Valkanov, R. (2006). Linear time series processes with mixed data sampling and MIDAS regression models. *Mimeo*.

Schumacher, C. (2014). MIDAS regressions with time-varying parameters: An application to corporate bond spreads and GDP in the euro area. Annual conference 2014 (Hamburg): Evidence-based economic policy, Verein für Socialpolitik / German Economic Association.

Appendix 1 - VAR(1) stock

The equations defining the DGP corresponding to a VAR(1) are the following:

$$y_t \ = \ \phi_l y_{t-1} \ + \ \delta_l x_{t-1} \ + \ e_{y,t} \tag{31a}$$

$$x_t \ = \ \phi_h x_{t-1} \ + \ e_{x,t} \tag{31b}$$

According to Equation 31a, we can write:

$$y_{t-k} \ = \ \phi_l y_{t-k-1} \ + \ \delta_l x_{t-k-1} \ + \ e_{y,t-k} \qquad\qquad t = 1, 2, 3, ... \tag{32}$$

Considering the previous expression for $k = 1$ and $k = 2$, and substituting them in Equation 31a, we obtain:

$$\begin{aligned} y_t \ &= \phi_l[\phi_l y_{t-2} \ + \ \delta_l x_{t-2} \ + \ e_{y,t-1}] \ + \ \delta_l x_{t-1} \ + \ e_{y,t} \\ &= \phi_l^2 y_{t-2} \ + \ \phi_l \delta_l x_{t-2} \ + \ \phi_l e_{y,t-1} \ + \ \delta_l x_{t-1} \ + \ e_{y,t} \\ &= \phi_l^2[\phi_l y_{t-3} \ + \ \delta_l x_{t-3} \ + \ e_{y,t-2}] \ + \ \phi_l \delta_l x_{t-2} \ + \ \phi_l e_{y,t-1} \ + \ \delta_l x_{t-1} \ + \ e_{y,t} \\ &= \phi_l^3 y_{t-3} \ + \ \phi_l^2 \delta_l x_{t-3} \ + \ \phi_l^2 e_{y,t-2} \ + \ \phi_l \delta_l x_{t-2} \ + \ \phi_l e_{y,t-1} \ + \ \delta_l x_{t-1} \ + \ e_{y,t} \end{aligned} \tag{33}$$

Rearranging terms, we obtain the following equation:

$$(1 \ - \ \phi_l^3 L^3) \ y_t \ = \ \delta_l x_{t-1} \ + \ \phi_l \delta_l x_{t-2} \ + \ \phi_l^2 \delta_l x_{t-3} \ + \ [1 + \phi_l L + \phi_l^2 L^2] \ e_{y,t} \tag{34}$$

We can then rewrite the previous equation in TF form, expressing the aggregated value of y (i.e., y_t^A) in terms of HF values of variable x and an error term:

$$\begin{aligned} y_t^A &= \frac{\delta_l}{1 - \phi_l^3 L^3} x_{t-1} + \frac{\phi_l \delta_l}{1 - \phi_l^3 L^3} x_{t-2} + \frac{\phi_l^2 \delta_l}{1 - \phi_l^3 L^3} x_{t-3} + \frac{1 + \phi_l L + \phi_l^2 L^2}{1 - \phi_l^3 L^3} e_{y,t} \\ &= \frac{\delta_l}{1 - \phi_l^3 L^3} x_{t-1} + \frac{\phi_l \delta_l}{1 - \phi_l^3 L^3} x_{t-2} + \frac{\phi_l^2 \delta_l}{1 - \phi_l^3 L^3} x_{t-3} + \frac{1}{1 - \phi_l L} e_{y,t} \qquad t = 4, 5, 6, ... \end{aligned} \tag{35}$$

where to obtain the last expression we have applied that $1 - \phi_l^3 L^3 = (1 - \phi_l L)(1 + \phi_l L + \phi_l^2 L^2)$.

Finally, if we consider the aggregated value of y for each quarter (i.e., $y_{t_q}^A$), from the first expression of the previous equation we can derive that:

$$y_{t_q}^A = \frac{\delta_l}{1 - \phi_l^3 L} x_{2,t_q} + \frac{\phi_l \delta_l}{1 - \phi_l^3 L} x_{1,t_q} + \frac{\phi_l^2 \delta_l}{1 - \phi_l^3 L} x_{3,t_q-1} + \frac{1}{1 - \phi_l^3 L} \eta_{\ t_q} \qquad t_q = 2, 3, ... \tag{36}$$

where $y_{t_q}^A$ is the aggregated quarterly value of y, which is a stock variable (i.e. $y_{t_q}^A = y_t$ for $t = 3 * t_q$); x_{2,t_q} is the second monthly value of x for current quarter; x_{1,t_q} is the first monthly value of x for current quarter; x_{3,t_q-1} is the last monthly value of x for the previous quarter; and $\eta_{\ t_q} = (1 + \phi_l L + \phi_l^2 L^2)\ e_{y,t}$.

Let's now analyse the structure of the error term $\eta_{\ y^A,t_q}$ in Equation 36.

$$E[\eta_{\ t_q}] = E[(1 + \phi_l L + \phi_l^2 L^2)\ e_{y,t}] = 0 \tag{37}$$

$$\begin{aligned} V[\eta_{\ t_q}] &= V[(1 + \phi_l L + \phi_l^2 L^2)\ e_{y,t}] = V[e_{y,t}] + \phi_l^2\ V[e_{y,t-1}] + \phi_l^4\ V[e_{y,t-2}] \\ &= (1 + \phi_l^2 + \phi_l^4)\ V[e_{y,t}] \end{aligned} \tag{38}$$

$$
\begin{aligned}
Cov(\eta_{\,t_q},\ \eta_{\,t_q-k}) &= E[\eta_{\,t_q}\ *\ \eta_{\,t_q-k}] \\
&= E[((1\ +\ \phi_l L\ +\ \phi_l^2 L^2)\ e_{y,t})\ *\ ((1\ +\ \phi_l L\ +\ \phi_l^2 L^2)\ e_{y,t-3k})] \\
&= E[(e_{y,t}\ +\ \phi_l e_{y,t-1}\ +\ \phi_l^2 e_{y,t-2}) * (e_{y,t-3k}\ +\ \phi_l e_{y,t-3k-1}\ +\ \phi_l^2 e_{y,t-3k-2})] \\
&= 0 \qquad\qquad\qquad k = 1, 2, 3, ...
\end{aligned} \tag{39}
$$

Therefore we can state that:

$$
\rho_k\ =\ 0 \qquad\qquad k = 1, 2, 3, ... \tag{40}
$$

The error term $\eta_{\,t_q}$ is then a white noise.

Finally, we can rewrite Equation 36 and express the aggregated value of a stock variable y for each quarter (i.e., $y^A_{t_q}$) in TF form:

$$
y^A_{t_q}\ =\ \frac{\delta_l}{1\ -\ \phi_l^3 L}\ x_{2,t_q}\ +\ \frac{\phi_l \delta_l}{1\ -\ \phi_l^3 L}\ x_{1,t_q}\ +\ \frac{\phi_l^2 \delta_l}{1\ -\ \phi_l^3 L}\ x_{3,t_q-1} +\ \zeta_{t_q} \tag{41a}
$$

$$
(1\ -\ \phi_l^3 L)\ \zeta_{t_q}\ =\ e_{y^A,t_q} \qquad\qquad t_q = 2, 3, ... \tag{41b}
$$

where the error term ζ_{t_q} has an AR(1) structure, and e_{y^A,t_q} is a white noise, uncorrelated with the explicative variables x_{1,t_q}, x_{2,t_q} and x_{3,t_q}, and with this variance:

$$
V[e_{y^A,t_q}]\ =\ V[\eta_{\,t_q}]\ =\ (1 + \phi_l^2 + \phi_l^4)\ V[e_{y,t}] \tag{42}
$$

Appendix 2 - VAR(1) flow

From the expression for y_t obtained in Equation 35 we can deduce equivalent expressions for y_{t-1} and y_{t-2}:

$$y_{t-1} = \frac{\delta_l}{1-\phi_l^3 L^3} x_{t-2} \;+\; \frac{\phi_l \delta_l}{1-\phi_l^3 L^3} x_{t-3} \;+\; \frac{\phi_l^2 \delta_l L^3}{1-\phi_l^3 L^3} x_{t-1} \;+\; \frac{L + \phi_l L^2 + \phi_l^2 L^3}{1-\phi_l^3 L^3} e_{y,t} \quad (43a)$$

$$y_{t-2} = \frac{\delta_l}{1-\phi_l^3 L^3} x_{t-3} \;+\; \frac{\phi_l \delta_l L^3}{1-\phi_l^3 L^3} x_{t-1} \;+\; \frac{\phi_l^2 \delta_l L^3}{1-\phi_l^3 L^3} x_{t-2} \;+\; \frac{L^2 + \phi_l L^3 + \phi_l^2 L^4}{1-\phi_l^3 L^3} e_{y,t} \quad (43b)$$

In order to arrive at the previous equations we have recalled that $x_{t-4} = L^3 x_{t-1}$, $x_{t-5} = L^3 x_{t-2}$, $e_{y,t-1} = L e_{y,t}$ and $e_{y,t-2} = L^2 e_{y,t}$.

Adding the previous expressions for y_t, y_{t-1} and y_{t-2} we obtain the aggregated quarterly value of y in the case of a flow variable (i.e., y_t^A):

$$\begin{aligned}
y_t^A &= y_t \;+\; y_{t-1} \;+\; y_{t-2} \\
&= [\frac{\delta_l}{1 \;-\; \phi_l^3 L^3}\; x_{t-1} \;+\; \frac{\phi_l \delta_l}{1 \;-\; \phi_l^3 L^3}\; x_{t-2} \;+\; \frac{\phi_l^2 \delta_l}{1 \;-\; \phi_l^3 L^3}\; x_{t-3} \;+\; \frac{1 + \phi_l L + \phi_l^2 L^2}{1 \;-\; \phi_l^3 L^3}\; e_{y,t}] \\
&\quad + [\frac{\delta_l}{1 \;-\; \phi_l^3 L^3}\; x_{t-2} \;+\; \frac{\phi_l \delta_l}{1 \;-\; \phi_l^3 L^3}\; x_{t-3} \;+\; \frac{\phi_l^2 \delta_l L^3}{1 \;-\; \phi_l^3 L^3}\; x_{t-1} \;+\; \frac{L + \phi_l L^2 + \phi_l^2 L^3}{1 \;-\; \phi_l^3 L^3}\; e_{y,t}] \\
&\quad + [\frac{\delta_l}{1 \;-\; \phi_l^3 L^3}\; x_{t-3} \;+\; \frac{\phi_l \delta_l L^3}{1 \;-\; \phi_l^3 L^3}\; x_{t-1} \;+\; \frac{\phi_l^2 \delta_l L^3}{1 \;-\; \phi_l^3 L^3}\; x_{t-2} \;+\; \frac{L^2 + \phi_l L^3 + \phi_l^2 L^4}{1 \;-\; \phi_l^3 L^3}\; e_{y,t}]
\end{aligned} \quad (44)$$

Re-organizing terms we obtain the expression in TF form of the quarterly aggregated value of y (i.e., y_t^A) in the case of a flow variable:

$$y_t^A \;=\; \frac{\delta_l \;+\; [\phi_l\delta_l(1+\phi_l)]L^3}{1 \;-\; \phi_l^3 L^3}\; x_{t-1} \;+\; \frac{\delta_l(1+\phi_l) \;+\; (\phi_l^2\delta_l)L^3}{1 \;-\; \phi_l^3 L^3}\; x_{t-2}$$

$$+\; \frac{\delta_l(1+\phi_l+\phi_l^2)}{1 \;-\; \phi_l^3 L^3}\; x_{t-3} \;+\; \frac{1 \;+\; (1+\phi_l)L \;+\; (1+\phi_l+\phi_l^2)L^2 \;+\; \phi_l(1+\phi_l)L^3 \;+\; \phi_l^2 L^4}{1 \;-\; \phi_l^3 L^3}\; e_{y,t}$$

$$=\; \frac{\delta_l \;+\; [\phi_l\delta_l(1+\phi_l)]L^3}{1 \;-\; \phi_l^3 L^3}\; x_{t-1} \;+\; \frac{\delta_l(1+\phi_l) \;+\; (\phi_l^2\delta_l)L^3}{1 \;-\; \phi_l^3 L^3}\; x_{t-2}$$

$$+\; \frac{\delta_l(1+\phi_l+\phi_l^2)}{1 \;-\; \phi_l^3 L^3}\; x_{t-3} \;+\; \frac{1 \;+\; L \;+\; L^2}{1 \;-\; \phi_l L}\; e_{y,t} \qquad t = 4, 5, 6, ... \tag{45}$$

Finally, from the first equality in the previous equation we can derive this result that expresses the LF aggregated value of a flow variable y (i.e., $y_{t_q}^A$) in terms of the HF values of x :

$$y_{t_q}^A \;=\; \frac{\delta_l \;+\; [\phi_l\delta_l(1+\phi_l)]L}{1 \;-\; \phi_l^3 L}\; x_{2,t_q} \;+\; \frac{\delta_l(1+\phi_l) \;+\; (\phi_l^2\delta_l)L}{1 \;-\; \phi_l^3 L}\; x_{1,t_q}$$

$$+\; \frac{\delta_l(1+\phi_l+\phi_l^2)}{1 \;-\; \phi_l^3 L}\; x_{3,t_q-1} \;+\; \frac{1}{1 \;-\; \phi_l^3 L}\; \eta_{\,t_q} \qquad t_q = 2, 3, 4, ... \tag{46}$$

where $y_{t_q}^A$ is the aggregated quarterly value of y, which is a flow variable (i.e. $y_{t_q}^A = y_t + y_{t-1} + y_{t-2}$ for $t = 3 * t_q$); x_{2,t_q} is the second monthly value of x for current quarter; x_{1,t_q} is the first monthly value of x for current quarter; x_{3,t_q-1} is the last monthly value of x for the previous quarter; and $\eta_{\,t_q} = [1 + (1+\phi_l)L + (1+\phi_l+\phi_l^2)L^2 + \phi_l(1+\phi_l)L^3 + \phi_l^2 L^4]\; e_{y,t}$.

Let's analyse the structure of the error term $\eta_{\,t_q}$ in Equation 46.

$$E[(1 \;+\; (1+\phi_l)L \;+\; (1+\phi_l+\phi_l^2)L^2 \;+\; \phi_l(1+\phi_l)L^3 \;+\; \phi_l^2 L^4)\; e_{y,t}] \;=\; 0 \tag{47}$$

$$V[(1\ +\ (1+\phi_l)L\ +\ (1+\phi_l+\phi_l^2)L^2\ +\ \phi_l(1+\phi_l)L^3\ +\ \phi_l^2L^4)\ e_{y,t}]$$

$$\begin{aligned} &= V[e_{y,t}]\ +\ (1+\phi_l)^2\ V[e_{y,t-1}] +\ (1+\phi_l+\phi_l^2)^2\ V[e_{y,t-2}] \\ &\quad +\ [\phi_l(1+\phi_l)]^2\ V[e_{y,t-3}]\ +\ \phi_l^4\ V[e_{y,t-4}] \end{aligned} \tag{48}$$

$$= (3\ +\ 4\phi_l\ +\ 5\phi_l^2\ +\ 4\phi_l^3\ +\ 3\phi_l^4)\ V[e_{y,t}]$$

$$Cov[(1\ +\ (1+\phi_l)L\ +\ (1+\phi_l+\phi_l^2)L^2\ +\ \phi_l(1+\phi_l)L^3\ +\ \phi_l^2L^4)\ e_{y,t},$$

$$(1\ +\ (1+\phi_l)L\ +\ (1+\phi_l+\phi_l^2)L^2\ +\ \phi_l(1+\phi_l)L^3\ +\ \phi_l^2L^4)\ e_{y,t-3}]$$

$$\begin{aligned} &= E[(1\ +\ (1+\phi_l)L\ +\ (1+\phi_l+\phi_l^2)L^2\ +\ \phi_l(1+\phi_l)L^3\ +\ \phi_l^2L^4)\ e_{y,t} \\ &\quad *\ (1\ +\ (1+\phi_l)L\ +\ (1+\phi_l+\phi_l^2)L^2\ +\ \phi_l(1+\phi_l)L^3\ +\ \phi_l^2L^4)\ e_{y,t-3}] \end{aligned} \tag{49a}$$

$$= \phi_l(1+\phi_l)\ V[e_{y,t-3}]\ +\ \phi_l^2(1+\phi_l)\ V[e_{y,t-4}]$$

$$= \phi_l(1+2\phi_l+\phi_l^2)\ V[e_{y,t}]$$

$$Cov[(1\ +\ (1+\phi_l)L\ +\ (1+\phi_l+\phi_l^2)L^2\ +\ \phi_l(1+\phi_l)L^3\ +\ \phi_l^2L^4)\ e_{y,t},$$

$$(1\ +\ (1+\phi_l)L\ +\ (1+\phi_l+\phi_l^2)L^2\ +\ \phi_l(1+\phi_l)L^3\ +\ \phi_l^2L^4)\ e_{y,t-3k}]\ =\ 0 \tag{49b}$$

$$k = 2, 3, 4, ...$$

Therefore we can state that:

$$\rho_1\ =\ \frac{\phi_l(1+2\phi_l+\phi_l^2)\ V[e_{y,t}]}{(3\ +\ 4\phi_l\ +\ 5\phi_l^2\ +\ 4\phi_l^3\ +\ 3\phi_l^4)\ V[e_{y,t}]}\ =\ \frac{\phi_l(1+2\phi_l+\phi_l^2)}{3\ +\ 4\phi_l\ +\ 5\phi_l^2\ +\ 4\phi_l^3\ +\ 3\phi_l^4} \tag{50a}$$

$$\rho_k\ =\ 0 \qquad\qquad k = 2, 3, 4, ... \tag{50b}$$

Then, as $\eta_{\ t_q}$ has a MA(1) structure, we can state that:

$$\eta_{\ t_q} \ = \ (1 + \psi L)\ e_{y^A, t_q} \tag{51}$$

where e_{y^A, t_q} is a white noise.

We can derive the value of ψ in Equation 51, recalling that the acf for a MA(1) model has the following expression:

$$\rho_1 \ = \ \frac{\psi}{1 \ + \ \psi^2} \tag{52a}$$

$$\rho_k \ = \ 0 \qquad\qquad k = 2, 3, ... \tag{52b}$$

From Equation 52a, we can derive the value of ψ:

$$\rho_1 \psi^2 \ - \ \psi \ + \ \rho_1 \ = \ 0 \tag{53}$$

Solving the previous equation we arrive at this result:

$$\psi \ = \ \frac{1 \ \pm \ \sqrt{1 \ - \ 4\rho_1^2}}{2\rho_1} \tag{54}$$

Between the two solutions for Equation 54, we consider the one which satisfies the invertibility condition: $|\psi| < 1$.

Therefore, we can finally state that the LF aggregated value of a flow variable y (i.e., $y^A_{t_q}$) can be expressed as a model in TF form in the following way:

$$y^A_{t_q} \;=\; \frac{\delta_l \;+\; [\phi_l \delta_l (1+\phi_l)]L}{1 \;-\; \phi_l^3 L}\; x_{2,t_q} \;+\; \frac{\delta_l(1+\phi_l) \;+\; (\phi_l^2 \delta_l)L}{1 \;-\; \phi_l^3 L}\; x_{1,t_q}$$

$$+\; \frac{\delta_l(1+\phi_l+\phi_l^2)}{1 \;-\; \phi_l^3 L}\; x_{3,t_q-1} \;+\; \zeta_{t_q} \qquad t_q = 2,3,4,... \tag{55a}$$

$$(1 \;-\; \phi_l^3 L)\; \zeta_{t_q} \;=\; (1 \;+\; \psi L)\; e_{y^A,t_q} \tag{55b}$$

where the error term ζ_{t_q} has an ARMA(1,1) structure, and e_{y^A,t_q} is a white noise, uncorrelated with the explicative variables x_{1,t_q}, x_{2,t_q} and x_{3,t_q}, with the folowing variance.

From Equations 48 and 51 we can derive an expression for $V[e_{y^A,t_q}]$:

$$\begin{aligned} V[\eta_{\;t_q}] \;&=\; V[(1+\psi L)\; e_{y^A,t_q}] \;=\; V[e_{y^A,t_q} \;+\; \psi\; e_{y^A,t_q-1}] \\ &=\; V[e_{y^A,t_q}] \;+\; \psi^2\; V[e_{y^A,t_q-1}] \;=\; (1+\psi^2)\; V[e_{y^A,t_q}] \end{aligned} \tag{56}$$

Therefore we can conclude that:

$$V[e_{y^A,t_q}] \;=\; \frac{V[\eta_{\;t_q}]}{1 \;+\; \psi^2} \;=\; \frac{(3 \;+\; 4\phi_l \;+\; 5\phi_l^2 \;+\; 4\phi_l^3 \;+\; 3\phi_l^4)}{1 \;+\; \psi^2}\; V[e_{y,t}] \tag{57}$$

Appendix 3 - VMA(1) stock

The equations defining the DGP corresponding to a VMA(1), where we assume again for simplicity that $\delta_h = 0$ [i.e., past values of y do not affect the present value of x], are the following:

$$y_t \;=\; e_{y,t} \;+\; \theta_l e_{y,t-1} \;+\; \delta_l e_{x,t-1} \tag{58a}$$

$$x_t \;=\; e_{x,t} \;+\; \theta_h e_{x,t-1} \tag{58b}$$

From Equation 58b we can state that:

$$x_{t-k} \;=\; e_{x,t-k} \;+\; \theta_h e_{x,t-k-1} \qquad\qquad k = 1, 2, 3, ... \tag{59}$$

And from this equation we can derive that:

$$e_{x,t-k} \;=\; x_{t-k} \;-\; \theta_h e_{x,t-k-1} \qquad\qquad k = 1, 2, ... \tag{60}$$

Taking into account Equation 60, we deduce that $e_{x,t-1} \;=\; x_{t-1} \;/\; (1 \;+\; \theta_h L)$. Re-arranging terms, we can rewrite Equation 58a as follows:

$$\begin{aligned} y_t \;&= e_{y,t} \;+\; \theta_l e_{y,t-1} \;+\; \delta_l \;[\frac{x_{t-1}}{1 \;+\; \theta_h L}] \\ &= \frac{\delta_l}{1 \;+\; \theta_h L}\; x_{t-1} \;+\; (1 + \theta_l L)\; e_{y,t} \end{aligned} \tag{61}$$

where we have applied that $(1 \;+\; \theta_h L)(1 \;+\; \theta_l L) \;=\; 1 \;+\; (\theta_h + \theta_l)\; L \;+\; (\theta_h \theta_l)\; L^2$.

If we develop the first term of the previous equation according a geometric progression with common ratio $1+\theta_h L,$ we can rewrite Equation 61 in the following way:

$$
\begin{aligned}
y_t^A \;&=\; \delta_l \sum_{i=0}^{\infty}(-\theta_h^3 L^3)^i \; x_{t-1} \;-\; \delta_l \theta_h \sum_{i=0}^{\infty}(-\theta_h^3 L^3)^i \; x_{t-2} \\
&\quad + \; \delta_l \theta_h^2 \sum_{i=0}^{\infty}(-\theta_h^3 L^3)^i \; x_{t-3} \;+\; (1 \;+\; \theta_l L) \; e_{y,t} \\
&= \frac{\delta_l}{1 \;+\; \theta_h^3 L^3} \, x_{t-1} \;-\; \frac{\delta_l \theta_h}{1 \;+\; \theta_h^3 L^3} \, x_{t-2} \;+\; \frac{\delta_l \theta_h^2}{1 \;+\; \theta_h^3 L^3} \, x_{t-3} \;+\; (1+\theta_l L) \; e_{y,t} \\
&\qquad\qquad t = 4, 5, ...
\end{aligned}
\tag{62}
$$

Considering the aggregated value of the stock variable y for each quarter (i.e., $y_{t_q}^A$), we can rewrite the previous equation in this way:

$$
y_{t_q}^A \;=\; \frac{\delta_l}{1 \;+\; \theta_h^3 L} \, x_{2,t_q} \;-\; \frac{\delta_l \theta_h}{1 \;+\; \theta_h^3 L} \, x_{1,t_q} \;+\; \frac{\delta_l \theta_h^2}{1 \;+\; \theta_h^3 L} \, x_{3,t_q-1} \;+\; \eta_{\,t_q} \qquad t_q = 2, 3, ...
\tag{63}
$$

where $y_{t_q}^A$ is the aggregated quarterly value of y, which is a stock variable (i.e. $y_{t_q}^A = y_t$ for $t = 3 * t_q$); x_{2,t_q} is the second monthly value of x for current quarter; x_{1,t_q} is the first monthly value of x for current quarter; x_{3,t_q-1} is the last monthly value of x for the previous quarter; and $\eta_{\,t_q} = (1+\theta_l L) \; e_{y,t}$.

Let's analyse the structure of the error term $\eta_{\,t_q}$:

$$
E[\eta_{\,t_q}] \;=\; E[(1 \;+\; \theta_l L) \; e_{y,t}] \;=\; 0
\tag{64}
$$

$$
V[\eta_{\,t_q}] \;=\; V[(1 \;+\; \theta_l L) \; e_{y,t}] \;=\; V[e_{y,t}] \;+\; \theta_l^2 \; V[e_{y,t-1}] \;=\; (1 \;+\; \theta_l^2) \; V[e_{y,t}]
\tag{65}
$$

$$Cov(\eta_{\,t_q},\ \eta_{\,t_q-k}) \ = \ E[\eta_{\,t_q} \ * \ \eta_{\,t_q-k}] \ = \ E[(1+\theta_l L)\ e_{y,t} \ * \ (1+\theta_l L)\ e_{y,t-3k})]$$

$$= \ E[(e_{y,t} \ + \ \theta_l e_{y,t-1}) \ * \ (e_{y,t-3k} \ + \ \theta_l e_{y,t-3k-1})] \ = \ 0 \qquad k=1,2,... \tag{66}$$

Thus we can state:

$$\rho_k \ = \ 0 \qquad k=1,2,3,... \tag{67}$$

Finally, we can conclude that the expression for the aggregated quarterly value of y in the case of a stock variable is:

$$y^A_{t_q} \ = \ \frac{\delta_l}{1 \ + \ \theta_h^3 L}\, x_{2,t_q} \ - \ \frac{\delta_l \theta_h}{1 \ + \ \theta_h^3 L}\, x_{1,t_q} \ + \ \frac{\delta_l \theta_h^2}{1 \ + \ \theta_h^3 L}\, x_{3,t_q-1} \ + \ \zeta_{t_q} \tag{68a}$$

$$\zeta_{t_q} \ = \ e_{y^A,t_q} \qquad t_q = 2,3,4,... \tag{68b}$$

where the error term ζ_{t_q} is a white noise, and e_{y^A,t_q} is a white noise, uncorrelated with the explicative variables x_{1,t_q}, x_{2,t_q} and x_{3,t_q}, with variance: $V[e_{y^A,t_q}] \ = \ V[\eta_{\,t_q}] \ = \ (1 \ + \ \theta_l^2)\ V[e_{y,t}]$.

Appendix 4 - VMA(1) flow

From the expression for y_t obtained in Equation 62 we can deduce equivalent expressions for y_{t-1} and y_{t-2}:

$$y_{t-1} = \frac{\delta_l}{1 + \theta_h^3 L^3} x_{t-2} - \frac{\delta_l \theta_h}{1 + \theta_h^3 L^3} x_{t-3} + \frac{\delta_l \theta_h^2 L^3}{1 + \theta_h^3 L^3} x_{t-1} + (L + \theta_l L^2) e_{y,t} \tag{69a}$$

$$y_{t-2} = \frac{\delta_l}{1 + \theta_h^3 L^3} x_{t-3} - \frac{\delta_l \theta_h L^3}{1 + \theta_h^3 L^3} x_{t-1} + \frac{\delta_l \theta_h^2 L^3}{1 + \theta_h^3 L^3} x_{t-2} + (L^2 + \theta_l L^3) e_{y,t} \tag{69b}$$

In order to arrive at the previous equations we have recalled that $x_{t-4} = L^3 x_{t-1}$, $x_{t-5} = L^3 x_{t-2}$, $e_{y,t-1} = L e_{y,t}$ and $e_{y,t-2} = L^2 e_{y,t}$.

Adding the previous expressions for y_t, y_{t-1} and y_{t-2} we obtain the aggregated value of y (i.e., y_t^A) in the case of a flow variable:

$$\begin{aligned}
y_t^A &= y_t + y_{t-1} + y_{t-2} \\
&= [\frac{\delta_l}{1 + \theta_h^3 L^3} x_{t-1} - \frac{\delta_l \theta_h}{1 + \theta_h^3 L^3} x_{t-2} + \frac{\delta_l \theta_h^2}{1 + \theta_h^3 L^3} x_{t-3} + (1 + \theta_l L) e_{y,t}] \\
&\quad + [\frac{\delta_l}{1 + \theta_h^3 L^3} x_{t-2} - \frac{\delta_l \theta_h}{1 + \theta_h^3 L^3} x_{t-3} + \frac{\delta_l \theta_h^2 L^3}{1 + \theta_h^3 L^3} x_{t-1} + (L + \theta_l L^2) e_{y,t}] \\
&\quad + [\frac{\delta_l}{1 + \theta_h^3 L^3} x_{t-3} - \frac{\delta_l \theta_h L^3}{1 + \theta_h^3 L^3} x_{t-1} + \frac{\delta_l \theta_h^2 L^3}{1 + \theta_h^3 L^3} x_{t-2} + (L^2 + \theta_l L^3) e_{y,t}]
\end{aligned} \tag{70}$$

Re-organizing terms we obtain the expression in TF form of the quarterly aggregated value of a flow variable y :

$$y_t^A = \frac{\delta_l + \delta_l\theta_h(\theta_h - 1)L^3}{1 + \theta_h^3 L^3} x_{t-1} + \frac{\delta_l(1-\theta_h) + \delta_l\theta_h^2 L^3}{1 + \theta_h^3 L^3} x_{t-2} + \frac{\delta_l(1-\theta_h+\theta_h^2)}{1 + \theta_h^3 L^3} x_{t-3}$$

$$+ \quad (1+\theta_l L)(1+L+L^2)\ e_{y,t} \qquad t = 4, 5, ... \tag{71}$$

where we have applied that $1+(1+\theta_l)L+(1+\theta_l)L^2+\theta_l L^3 = (1+\theta_l L)(1+L+L^2)$.

If we consider the aggregated value of the flow variable y for each quarter (i.e., $y_{t_q}^A$), we can rewrite the previous equation in this way:

$$\begin{aligned} y_{t_q}^A = & \frac{\delta_l + \delta_l\theta_h(\theta_h - 1)L}{1 + \theta_h^3 L} x_{2,t_q} + \frac{\delta_l(1-\theta_h) + \delta_l\theta_h^2 L}{1 + \theta_h^3 L} x_{1,t_q} \\ & + \frac{\delta_l(1-\theta_h+\theta_h^2)}{1 + \theta_h^3 L} x_{3,t_q-1} + \eta_{\ t_q} \qquad t_q = 2, 3, ... \end{aligned} \tag{72}$$

where $y_{t_q}^A$ is the aggregated quarterly value of y, which is a stock variable (i.e. $y_{t_q}^A = y_t$ for $t = 3 * t_q$); x_{2,t_q} is the second monthly value of x for current quarter; x_{1,t_q} is the first monthly value of x for current quarter; x_{3,t_q-1} is the last monthly value of x for the previous quarter; and $\eta_{\ t_q} = [1 + (1+\theta_l)L + (1+\theta_l)L^2 + \theta_l L^3]\ e_{y,t}$.

Let's analyse the structure of the error term $\eta_{\ t_q}$.

$$E[\eta_{\ t_q}] = E\ [(1 + (1+\theta_l)L + (1+\theta_l)L^2 + \theta_l L^3)\ e_{y,t}] = 0 \tag{73}$$

$$\begin{aligned} V[\eta_{\ t_q}] &= V\ [(1 + (1+\theta_l)L + (1+\theta_l)L^2 + \theta_l L^3)\ e_{y,t}] \\ &= V[e_{y,t}] + (1+\theta_l)^2\ V[e_{y,t-1}] + (1+\theta_l)^2\ V[e_{y,t-2}] + \theta_l^2\ V[e_{y,t-3}] \\ &= [3 + 4\theta_l + 3\theta_l^2]\ V[e_{y,t}] \end{aligned} \tag{74}$$

$$Cov(\eta_{\ t_q},\ \eta_{\ t_q-1})\ =\ E[\eta_{\ t_q}\ *\ \eta_{\ t_q-1}]$$

$$= E[(1+(1+\theta_l)L+(1+\theta_l)L^2+\theta_l L^3)\ e_{y,t}\ *\ (1+(1+\theta_l)L+(1+\theta_l)L^2+\theta_l L^3)\ e_{y,t-3}]$$

$$= E[(e_{y,t}\ +\ (1+\theta_l)\ e_{y,t-1}\ +\ (1+\theta_l)\ e_{y,t-2}\ +\ \theta_l\ e_{y,t-3})$$

$$*\ (e_{y,t-3}\ +\ (1+\theta_l)\ e_{y,t-4}\ +\ (1+\theta_l)\ e_{y,t-5}\ +\ \theta_l\ e_{y,t-6})]$$

$$= \theta_l\ V[e_{y,t-3}] \tag{75a}$$

$$Cov(\eta_{\ t_q},\ \eta_{\ t_q-k})\ = E[\eta_{\ t_q}\ *\ \eta_{\ t_q-k}]$$

$$= E[(e_{y,t}\ +\ (1+\theta_l)\ e_{y,t-1}\ +\ (1+\theta_l)\ e_{y,t-2}\ +\ \theta_l\ e_{y,t-3})$$

$$*\ (e_{y,t-3k}\ +\ (1+\theta_l)\ e_{y,t-3k-1}\ +\ (1+\theta_l)\ e_{y,t-3k-2}\ +\ \theta_l\ e_{y,t-3k-3})]$$

$$=\ 0 \qquad\qquad k=2,3,... \tag{75b}$$

From the previous equations we can state that:

$$\rho_1\ =\ \frac{\theta_l\ V[e_{y,t}]}{(3\ +\ 4\theta_l\ +\ 3\theta_l^2)\ V[e_{y,t}]}\ =\ \frac{\theta_l}{3\ +\ 4\theta_l\ +\ 3\theta_l^2} \tag{76a}$$

$$\rho_k\ =\ 0 \qquad\qquad k=1,2,3,... \tag{76b}$$

Therefore we can deduce that the error term $\eta_{\ t_q}$ has a MA(1) structure.

Finally, we can conclude that the expression for the aggregated quarterly value of y in the case of a flow variable is:

$$
\begin{aligned}
y^A_{t_q} \;=\; & \frac{\delta_l \;+\; \delta_l \theta_h (\theta_h - 1) L}{1 \;+\; \theta_h^3 L} \; x_{2,t_q} \;+\; \frac{\delta_l (1 - \theta_h) \;+\; \delta_l \theta_h^2 L}{1 \;+\; \theta_h^3 L} \; x_{1,t_q} \\
& + \; \frac{\delta_l (1 - \theta_h + \theta_h^2)}{1 \;+\; \theta_h^3 L} \; x_{3,t_q - 1} \;+\; \zeta_{t_q}
\end{aligned}
\tag{77a}
$$

$$
\zeta_{t_q} \;= (1 \;+\; \psi L) \; e_{y^A, t_q} \qquad\qquad t_q = 2, 3, 4, ... \tag{77b}
$$

where the error term ζ_{t_q} has a MA(1) structure, and e_{y^A, t_q} is a white noise, uncorrelated with the explicative variables x_{1,t_q}, x_{2,t_q} and x_{3,t_q}. Coefficient ψ satisfies this condition:

$$
\psi \;=\; \frac{1 \;\pm\; \sqrt{1 \;-\; 4\rho_1^2}}{2\rho_1} \tag{78}
$$

Between the two solutions for Equation 78, we consider the one which satisfies the invertibility condition: $|\psi| < 1$.

The variance of e_{y^A, t_q} can be derived from Equations 74 and **??**, considering that $\eta_{\ t_q} \;=\; \zeta_{t_q}$:

$$
\begin{aligned}
V[\eta_{\ t_q}] \;&=\; V[(1 \;+\; \psi L) \; e_{y^A, t_q}] \;=\; V[e_{y^A, t_q} \;+\; \psi \; e_{y^A, t_q - 1}] \\
&=\; V[e_{y^A, t_q}] \;+\; \psi^2 \; V[e_{y^A, t_q - 1}] \;=\; (1 + \psi^2) \; V[e_{y^A, t_q}]
\end{aligned}
\tag{79}
$$

Therefore we can conclude that:

$$
V[e_{y^A, t_q}] \;=\; \frac{V[\eta_{\ t_q}]}{1 \;+\; \psi^2} \;=\; \frac{3 \;+\; 4\theta_l \;+\; 3\theta_l^2}{1 \;+\; \psi^2} \; V[e_{y,t}] \tag{80}
$$

Appendix 5 - VMA(3) stock

The equations defining the DGP corresponding to a VMA(3) are the following:

$$y_t = e_{y,t} + \theta_l\, e_{y,t-3} + \delta_l\, e_{x,t-1} \tag{81a}$$

$$x_t = e_{x,t} + \theta_h\, e_{x,t-1} \tag{81b}$$

From Equation 81a we can state that:

$$y_{t-3} = e_{y,t-3} + \theta_l\, e_{y,t-6} + \delta_l\, e_{x,t-4} \tag{82}$$

And from the previous equation we can derive this result:

$$e_{y,t-3} = y_{t-3} - \theta_l\, e_{y,t-6} - \delta_l\, e_{x,t-4} \tag{83}$$

Substituting the previous expression in Equation 81a we obtain:

$$\begin{aligned} y_t &= e_{y,t} + \theta_l\, y_{t-3} - \theta_l^2\, e_{y,t-6} - \theta_l \delta_l\, e_{x,t-4} + \delta_l\, e_{x,t-1} \\ &= e_{y,t} + \theta_l\, y_{t-3} - \theta_l^2\, e_{y,t-6} + \delta_l(1 - \theta_l L^3)\, e_{x,t-1} \end{aligned} \tag{84}$$

From Equation 81b we can derive that:

$$e_{x,t-1} = x_{t-1} - \theta_h\, e_{x,t-2} \qquad k = 1, 2, 3, ... \tag{85}$$

Substituting the previous result in Equation 84 we obtain this:

$$
\begin{aligned}
y_t &= e_{y,t} + \theta_l\, y_{t-3} - \theta_l^2\, e_{y,t-6} + \delta_l(1- \theta_l L^3)\; e_{x,t-1} \\
&= \theta_l\, y_{t-3} + (1- \theta_l^2 L^6)\, e_{y,t} + \delta_l(1- \theta_l L^3)\, [\frac{x_{t-1}}{1 + \theta_h L}]
\end{aligned}
\tag{86}
$$

We can rewrite Equation 86 in the following equivalent way:

$$
\begin{aligned}
y_t = \theta_l\, y_{t-3} &+ \frac{\delta_l(1 - \theta_l L^3)}{1 + \theta_h^3 L^3}\, x_{t-1} - \frac{\delta_l\theta_h(1 - \theta_l L^3)}{1 + \theta_h^3 L^3}\, x_{t-2} + \frac{\delta_l\theta_h^2(1 - \theta_l L^3)}{1 + \theta_h^3 L^3}\, x_{t-3} \\
&+ \frac{1 + \theta_h L - \theta_l^2 L^6 - \theta_h\theta_l^2 L^7}{1 + \theta_h L}\, e_{y,t} \qquad t = 4, 5, ...
\end{aligned}
\tag{87}
$$

We can then express the equation in TF form:

$$
y_t = \frac{\delta_l}{1 + \theta_h^3 L^3}\, x_{t-1} - \frac{\delta_l\theta_h}{1 + \theta_h^3 L^3}\, x_{t-2} + \frac{\delta_l\theta_h^2}{1 + \theta_h^3 L^3}\, x_{t-3} + (1+\theta_l L^3)\, e_{y,t} \qquad t = 4, 5, ...
\tag{88}
$$

where we have applied that: $(1-\theta_l L^3)(1+\theta_h L)(1+\theta_l L^3) = 1+\theta_h L-\theta_l^2 L^6-\theta_h\theta_l^2 L^7$.

Considering the aggregated value of a stock variable y for each quarter (i.e., $y_{t_q}^A$), we can re-write the previous equation in this way:

$$
y_{t_q}^A = \frac{\delta_l}{1 + \theta_h^3 L}\, x_{2,t_q} - \frac{\delta_l\theta_h}{1 + \theta_h^3 L}\, x_{1,t_q} + \frac{\delta_l\theta_h^2}{1 + \theta_h^3 L} x_{3,t_q-1} + (1+\theta_l L)\, e_{y^A,t_q} \qquad t_q = 2, 3, ...
\tag{89}
$$

where $y_{t_q}^A$ is the aggregated quarterly value of y, which is a stock variable (i.e. $y_{t_q}^A = y_t$ for $t = 3 * t_q$); x_{2,t_q} is the second monthly value of x for current quarter; x_{1,t_q} is the first monthly value of x for current quarter; x_{3,t_q-1} is the last monthly value of x for the previous quarter; and e_{y^A,t_q} is a white noise.

Therefore, we can finally state that the LF aggregated value of a stock variable y (i.e., $y^A_{t_q}$) can be expressed as a model in TF form in the following way:

$$y^A_{t_q} = \frac{\delta_l}{1 + \theta_h^3 L} x_{2,t_q} - \frac{\delta_l \theta_h}{1 + \theta_h^3 L} x_{1,t_q} + \frac{\delta_l \theta_h^2}{1 + \theta_h^3 L} x_{3,t_q-1} + \zeta_{t_q} \qquad t_q = 2, 3, ... \tag{90a}$$

$$\zeta_{t_q} = (1 + \theta_l L)\, e_{y^A,t_q} \tag{90b}$$

where the error term ζ_{t_q} has a MA(1) structure, and e_{y^A,t_q} is a white noise, uncorrelated with the explicative variables x_{1,t_q}, x_{2,t_q} and x_{3,t_q}, and with variance $V[e_{y^A,t_q}] = V[e_{y,t}]$.

Appendix 6 - VMA(3) flow

From the expression for y_t obtained in Equation 88 we can derive equivalent expressions for y_{t-1} and y_{t-2}:

$$y_{t-1} \;=\; \frac{\delta_l}{1 \;+\; \theta_h^3 L^3}\; x_{t-2} \;-\; \frac{\delta_l \theta_h}{1 \;+\; \theta_h^3 L^3}\; x_{t-3} \;+\; \frac{\delta_l \theta_h^2 \; L^3}{1 \;+\; \theta_h^3 L^3}\; x_{t-1} \;+\; (L \;+\; \theta_l L^4)\; e_{y,t} \tag{91a}$$

$$y_{t-2} \;=\; \frac{\delta_l}{1 \;+\; \theta_h^3 L^3}\; x_{t-3} \;-\; \frac{\delta_l \theta_h L^3}{1 \;+\; \theta_h^3 \; L^3}\; x_{t-1} \;+\; \frac{\delta_l \theta_h^2 \; L^3}{1 \;+\; \theta_h^3 L^3}\; x_{t-2} \;+\; (L^2 \;+\; \theta_l L^5)\; e_{y,t} \tag{91b}$$

In order to arrive at the previous equations we have recalled that $x_{t-4} = L^3 x_{t-1}$, $x_{t-5} = L^3 x_{t-2}$, $e_{y,t-1} = L e_{y,t}$ and $e_{y,t-2} = L^2 e_{y,t}$.

Adding the previous expressions for y_t, y_{t-1} and y_{t-2} we obtain the aggregated value of y (i.e., y_t^A):

$$\begin{aligned} y_t^A \; &= y_t \;+\; y_{t-1} \;+\; y_{t-2} \\ &= [\frac{\delta_l}{1 \;+\; \theta_h^3 L^3}\; x_{t-1} \;-\; \frac{\delta_l \theta_h}{1 \;+\; \theta_h^3 L^3}\; x_{t-2} \;+\; \frac{\delta_l \theta_h^2}{1 \;+\; \theta_h^3 L^3}\; x_{t-3} \;+\; (1 \;+\; \theta_l L^3)\; e_{y,t}] \\ &\quad +\; [\frac{\delta_l}{1 \;+\; \theta_h^3 L^3}\; x_{t-2} \;-\; \frac{\delta_l \theta_h}{1 \;+\; \theta_h^3 \; L^3}\; x_{t-3} \;+\; \frac{\delta_l \theta_h^2 L^3}{1 \;+\; \theta_h^3 L^3}\; x_{t-1} \;+\; (L \;+\; \theta_l L^4)\; e_{y,t}] \\ &\quad +\; [\frac{\delta_l}{1 \;+\; \theta_h^3 L^3}\; x_{t-3} \;-\; \frac{\delta_l \theta_h L^3}{1 \;+\; \theta_h^3 L^3}\; x_{t-1} \;+\; \frac{\delta_l \theta_h^2 L^3}{1 \;+\; \theta_h^3 L^3}\; x_{t-2} \;+\; (L^2 \;+\; \theta_l L^5)\; e_{y,t}] \end{aligned} \tag{92}$$

Re-organizing terms we obtain the expression in TF form of the quarterly aggregated value of y in the case of a flow variable:

$$y_t^A = \frac{\delta_l + \delta_l\theta_h(\theta_h - 1)L^3}{1 + \theta_h^3 L^3}\, x_{t-1} + \frac{\delta_l(1-\theta_h) + \delta_l\theta_h^2 L^3}{1 + \theta_h^3 L^3}\, x_{t-2} + \frac{\delta_l(1-\theta_h+\theta_h^2)}{1 + \theta_h^3 L^3}\, x_{t-3}$$

$$+ (1 + \theta_l L^3)(1 + L + L^2)\, e_{y,t} \qquad t = 4, 5, ... \tag{93}$$

Finally, if we consider the aggregated value of y for each quarter (i.e., $y_{t_q}^A$), we can re-write the previous equation in this way:

$$y_{t_q}^A = \frac{\delta_l + \delta_l\theta_h(\theta_h - 1)L}{1 + \theta_h^3 L}\, x_{t-1} - \frac{\delta_l(1-\theta_h) + \delta_l\theta_h^2 L}{1 + \theta_h^3 L}\, x_{t-2}$$
$$+ \frac{\delta_l(1-\theta_h+\theta_h^2)}{1 + \theta_h^3 L}\, x_{t-3} + (1 + \theta_l L)\, \eta_{\ t_q} \qquad t = 2, 3, ... \tag{94}$$

where $y_{t_q}^A$ is the aggregated quarterly value of y, which is a flow variable (i.e., $y_{t_q}^A = y_t + y_{t-1} + y_{t-2}$ for $t = 3 * t_q$); x_{2,t_q} is the second monthly value of x for current quarter; x_{1,t_q} is the first monthly value of x for current quarter; x_{3,t_q-1} is the last monthly value of x for the previous quarter; and $\eta_{\ t_q} = (1 + L + L^2)\ e_{y,t}$.

We analyse in the following lines the structure of the error term $\eta_{\ t_q}$.

$$E[\eta_{\ t_q}] = E[(1 + L + L^2)\ e_{y,t}] = 0 \tag{95}$$

$$V[\eta_{\ t_q}] = V[(1 + L + L^2)\ e_{y,t}] = 3\ V[e_{y,t}] \tag{96}$$

$$Cov(\eta_{\ t_q},\ \eta_{\ t_q-k}) \ = \ E[\eta_{\ t_q} \ * \ \eta_{\ t_q-1}] \ = \ E[(1+L+L^2)\ e_{y,t} \ * \ (1+L+L^2)\ e_{y,t-3k}]$$

$$= E[(e_{y,t}+e_{y,t-1}+e_{y,t-2}) \ * \ (e_{y,t-3k}+e_{y,t-3k-1}+e_{y,t-3k-2})] \ = \ 0$$

$$k = 1,2,3,... \tag{97}$$

Then we can derive that:

$$\rho_k = 0 \qquad\qquad k = 1,2,3,... \tag{98}$$

Therefore we can conclude that the expression for the aggregated quarterly value of y in the case of a flow variable is:

$$y^A_{t_q} = \frac{\delta_l \ + \ \delta_l\theta_h(\theta_h-1)L}{1 \ + \ \theta_h^3 L}\ x_{2,t_q} \ + \frac{\delta_l(1-\theta_h) \ + \ \delta_l\theta_h^2 L}{1 \ + \ \theta_h^3 L}\ x_{3,t_q} \ + \frac{\delta_l(1-\theta_h+\theta_h^2)}{1 \ + \ \theta_h^3 L} x_{1,t_q-1} \ + \zeta_{t_q} \tag{99a}$$

$$\zeta_{t_q} \ = \ (1 \ + \ \theta_l L)\ e_{y^A,t_q} \qquad\qquad t_q = 2,3,4,... \tag{99b}$$

where the error term ζ_{t_q} has a MA(1) structure, and e_{y^A,t_q} is a white noise, uncorrelated with the explicative variables x_{1,t_q}, x_{2,t_q} and x_{3,t_q}, with variance:

$$V[e_{y^A,t_q}] \ = \ V[\eta_{\ t_q}] \ = \ 3\ V[e_{y,t}] \tag{100}$$

Appendix 7

Table 5: Out-of-sample MSFE of TF-MIDAS relative to U-MIDAS (DGP: HF-VMA(1); aggregation method: skip-sampling)

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0	-0.5	0.1	0	0.987	0.947	0.968	0.986	1.007	1.030	0.983	0.945	0.965	0.983	1.004	1.024
0	-0.5	0.5	0	0.986	0.944	0.969	0.986	1.004	1.028	0.985	0.939	0.967	0.985	1.004	1.031
0	-0.5	1	0	1.041	0.943	0.977	1.005	1.078	1.188	1.038	0.934	0.973	1.004	1.080	1.190
0	-0.7	0.1	0	0.999	.948	0.967	0.985	1.008	1.037	0.998	0.946	0.967	0.984	1.007	1.037
0	-0.7	0.5	0	1.012	0.962	0.981	1.003	1.037	1.078	0.986	0.936	0.961	0.984	1.009	1.033
0	-0.7	1	0	1.036	0.940	0.980	1.015	1.071	1.154	1.022	0.913	0.964	1.000	1.066	1.147
0	-0.9	0.1	0	0.990	0.955	0.969	0.987	1.006	1.033	0.989	0.952	0.969	0.986	1.005	1.034
0	-0.9	0.5	0	0.987	0.938	0.963	0.985	1.009	1.038	0.985	0.925	0.958	0.982	1.008	1.042
0	-0.9	1	0	0.983	0.849	0.927	0.987	1.047	1.100	0.997	0.851	0.917	0.986	1.062	1.147
-0.5	0	0.1	0	0.991	0.955	0.971	0.986	1.010	1.039	0.988	0.949	0.969	0.985	1.007	1.031
-0.5	0	0.5	0	0.983	0.940	0.966	0.984	1.000	1.023	0.982	0.938	0.966	0.983	0.999	1.022
-0.5	0	1	0	0.982	0.935	0.965	0.984	1.001	1.028	0.982	0.935	0.963	0.984	1.000	1.028
-0.7	0	0.1	0	0.981	0.940	0.964	0.984	0.998	1.018	0.981	0.940	0.963	0.983	0.999	1.018
-0.7	0	0.5	0	0.990	0.942	0.969	0.988	1.005	1.038	0.986	0.939	0.964	0.986	1.002	1.028
-0.7	0	1	0	0.984	0.944	0.965	0.983	1.001	1.023	0.982	0.941	0.964	0.983	1.002	1.023
-0.9	0	0.1	0	0.983	0.942	0.965	0.982	0.998	1.026	0.982	0.941	0.963	0.982	0.998	1.027
-0.9	0	0.5	0	0.981	0.940	0.963	0.983	0.998	1.024	0.980	0.935	0.961	0.981	0.996	1.023
-0.9	0	1	0	0.988	0.945	0.967	0.985	1.005	1.030	0.983	0.943	0.964	0.983	1.002	1.023
-0.5	-0.5	0.1	0	0.988	0.947	0.968	0.985	1.006	1.037	0.982	0.940	0.964	0.981	1.001	1.027
-0.5	-0.5	0.5	0	0.985	0.948	0.967	0.984	1.001	1.025	0.982	0.942	0.961	0.983	1.002	1.024
-0.5	-0.5	1	0	1.027	0.939	0.972	1.000	1.070	1.150	1.025	0.932	0.966	1.001	1.071	1.151
-0.7	-0.7	0.1	0	0.985	0.946	0.969	0.984	1.002	1.030	0.984	0.942	0.967	0.983	1.000	1.027
-0.7	-0.7	0.5	0	1.004	0.954	0.974	0.995	1.029	1.071	0.982	0.936	0.960	0.978	1.003	1.032
-0.7	-0.7	1	0	1.018	0.939	0.963	0.996	1.045	1.116	1.018	0.923	0.958	0.995	1.055	1.139
-0.9	-0.9	0.1	0	0.984	0.946	0.966	0.984	1.001	1.026	0.983	0.942	0.963	0.982	1.001	1.027
-0.9	-0.9	0.5	0	0.985	0.940	0.963	0.985	1.007	1.034	0.981	0.928	0.959	0.982	1.004	1.035
-0.9	-0.9	1	0	1.016	0.937	0.975	1.007	1.055	1.096	0.993	0.901	0.946	0.985	1.035	1.083

In bold rows where median MSFE of TF-MIDAS is at least 5% better than U-MIDAS.

Table 6: Out-of-sample MSFE of TF-MIDAS relative to U-MIDAS (DGP: HF-VMA(1); aggregation method: addition)

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0	-0.5	0.1	0	0.996	0.953	0.973	0.989	1.011	1.048	0.988	0.943	0.967	0.986	1.005	1.034
0	-0.5	0.5	0	0.970	0.876	0.923	0.971	1.017	1.068	0.978	0.909	0.941	0.976	1.011	1.055
0	-0.5	1	0	0.886	0.746	0.817	**0.883**	0.958	1.019	0.983	0.921	0.953	0.981	1.011	1.047
0	-0.7	0.1	0	0.992	0.941	0.968	0.993	1.021	1.044	0.992	0.937	0.966	0.993	1.020	1.048
0	-0.7	0.5	0	1.004	0.936	0.972	0.999	1.038	1.103	1.029	0.931	0.971	1.011	1.088	1.160
0	-0.7	1	0	0.798	0.635	0.695	**0.781**	0.904	1.013	1.010	0.880	0.930	0.983	1.054	1.181
0	-0.9	0.1	0	0.997	0.948	0.970	0.994	1.018	1.051	0.996	0.943	0.970	0.993	1.018	1.054
0	-0.9	0.5	0	0.932	0.794	0.860	**0.938**	1.000	1.067	0.962	0.859	0.910	0.962	1.011	1.065
0	-0.9	1	0	0.750	0.547	0.630	**0.739**	0.852	0.961	1.008	0.785	0.869	0.964	1.097	1.295
-0.5	0	0.1	0	0.941	0.851	0.886	**0.941**	0.990	1.031	0.940	0.852	0.891	**0.941**	0.987	1.024
-0.5	0	0.5	0	0.957	0.876	0.918	0.954	0.991	1.038	0.953	0.870	0.914	0.953	0.988	1.033
-0.5	0	1	0	0.937	0.846	0.891	**0.938**	0.978	1.035	0.932	0.852	0.889	**0.930**	0.970	1.019
-0.7	0	0.1	0	0.943	0.849	0.889	**0.937**	0.994	1.045	0.941	0.848	0.885	**0.935**	0.992	1.044
-0.7	0	0.5	0	0.880	0.779	0.820	**0.869**	0.934	0.990	0.879	0.780	0.821	**0.870**	0.931	0.988
-0.7	0	1	0	0.830	0.718	0.763	**0.822**	0.886	0.955	0.884	0.766	0.812	**0.877**	0.939	1.004
-0.9	0	0.1	0	0.806	0.693	0.741	**0.805**	0.875	0.990	0.805	0.692	0.740	**0.805**	0.871	0.982
-0.9	0	0.5	0	0.812	0.633	0.695	**0.779**	0.915	1.032	0.830	0.653	0.712	**0.801**	0.932	1.052
-0.9	0	1	0	0.735	0.586	0.631	**0.708**	0.825	0.936	0.829	0.666	0.718	**0.796**	0.910	1.050
-0.5	-0.5	0.1	0	0.997	0.940	0.963	0.993	1.020	1.068	0.989	0.933	0.957	0.988	1.012	1.057
-0.5	-0.5	0.5	0	0.882	0.761	0.812	**0.881**	0.937	1.005	0.935	0.853	0.894	**0.935**	0.973	1.015
-0.5	-0.5	1	0	0.731	0.580	0.647	**0.731**	0.803	0.875	0.958	0.854	0.901	**0.950**	1.009	1.067
-0.7	-0.7	0.1	0	0.956	0.863	0.906	0.957	0.998	1.051	0.955	0.863	0.906	0.956	0.998	1.049
-0.7	-0.7	0.5	0	0.752	0.602	0.676	**0.751**	0.827	0.897	0.910	0,785	0.840	**0.904**	0.969	1.044
-0.7	-0.7	1	0	0.466	0.361	0.401	**0.453**	0.521	0.598	0.864	0.739	0.796	**0.855**	0.923	0.989
-0.9	-0.9	0.1	0	0.838	0.723	0.772	**0.832**	0.889	0.990	0.837	0.724	0.771	**0.831**	0.889	0.988
-0.9	-0.9	0.5	0	0.635	0.462	0.526	**0.612**	0.733	0.835	0.771	0.566	0.642	**0.753**	0.885	0.997
-0.9	-0.9	1	0	0.480	0.291	0.359	**0.439**	0.586	0.733	0.879	0.573	0.644	**0.780**	1.093	1.329

In bold rows where median MSFE of TF-MIDAS is at least 5% better than U-MIDAS.

Table 7: Out-of-sample MSFE of TF-MIDAS relative to U-MIDAS (DGP: HF-VMA(3); aggregation method: skip-sampling)

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*	*Mean*	*10th percentile*	*25th percentile*	*Median*	*75th percentile*	*90th percentile*
0	-0.5	0.1	0	0.998	0.970	0.985	0.994	1.006	1.025	0.997	0.969	0.983	0.993	1.006	1.025
0	-0.5	0.5	0	0.996	0.958	0.978	0.995	1.013	1.037	0.993	0.951	0.973	0.992	1.013	1.040
0	-0.5	1	0	1.034	0.942	0.971	0.998	1.064	1.187	1.031	0.936	0.965	0.995	1.060	1.189
0	-0.7	0.1	0	0.987	0.944	0.965	0.985	1.004	1.031	0.986	0.945	0.963	0.985	1.04	1.033
0	-0.7	0.5	0	1.000	0.960	0.980	0.997	1.015	1.041	0.960	0.642	0760	**0.926**	1.123	1.324
0	-0.7	1	0	1.010	0.939	0.975	1.004	1.048	1.093	1.069	0.702	0.834	1.020	1.237	1.532
0	-0.9	0.1	0	0.985	0.944	0.967	0.985	1.005	1.041	0.984	0.943	0.966	0.984	1.011	1.051
0	-0.9	0.5	0	0.989	0.939	0.963	0.986	1.009	1.044	0.988	0.927	0.958	0.984	1.011	1.051
0	-0.9	1	0	0.941	0.814	0.867	**0.941**	1.005	1.060	0.940	0.819	0.872	**0.930**	1.011	1.069
-0.5	0	0.1	0	0.954	0.874	0.909	**0.950**	0.991	1.038	0.953	0.871	0.908	**0.950**	0.992	1.034
-0.5	0	0.5	0	0.924	0.828	0.872	**0.919**	0.975	1.022	0.922	0.827	0.869	**0.914**	0.974	1.022
-0.5	0	1	0	0.985	0.838	0.897	0.971	1.055	1.147	1.073	0.876	0.952	**1.052**	1.168	1.305
-0.7	0	0.1	0	0.950	0.864	0.905	**0.949**	0.993	1.037	0.950	0.861	0.904	**0.950**	0.992	1.038
-0.7	0	0.5	0	0.776	0.656	0.710	**0.768**	0.830	0.899	0.806	0.696	0.740	**0.796**	0.863	0.933
-0.7	0	1	0	0.774	0.667	0.706	**0.770**	0.838	0.922	0.808	0.694	0.739	**0.806**	0.872	0.960
-0.9	0	0.1	0												
-0.9	0	0.5	0	0.828	0.725	0.769	**0.816**	0.889	0.965	0.827	0.725	0.767	**0.818**	0.885	0.954
-0.9	0	1	0	0.699	0.577	0.624	**0.690**	0.763	0.843	0.889	0.552	0.685	**0.843**	1.043	1.247
-0.5	-0.5	0.1	0	0.956	0.876	0.910	0.952	0.998	1.044	0.955	0..874	0.912	0.952	0.999	1.043
-0.5	-0.5	0.5	0	0.927	0.829	0.877	**0.920**	0.973	1.028	0.926	0.838	0.876	**0.913**	0.974	1.029
-0.5	-0.5	1	0	0.874	0.770	0.818	**0.866**	0.922	0.980	0.912	0.799	0.844	**0.907**	0.970	1.025
-0.7	-0.7	0.1	0	0.873	0.767	0.812	**0.861**	0.926	1.000	0.871	0.766	0.808	**0.859**	0.926	1.000
-0.7	-0.7	0.5	0	0.849	0.731	0.780	**0.838**	0.912	0.973	0.851	0.735	0.787	**0.840**	0.908	0.976
-0.7	-0.7	1	0	0.762	0.639	0.693	**0.756**	0.829	0.898	0.848	0.707	0.773	**0.841**	0.920	1.000
-0.9	-0.9	0.1	0												
-0.9	-0.9	0.5	0	0.870	0.727	0.798	**0.868**	0.938	1.005	0.873	0.739	0.805	**0.872**	0.937	0.989
-0.9	-0.9	1	0	0.782	0.658	0.710	**0.770**	0.851	0.912	0.733	0.488	0.583	**0.707**	0.856	1.011

In bold rows where median MSFE of TF-MIDAS is at least 5% better than U-MIDAS.

Table 8: Out-of-sample MSFE of TF-MIDAS relative to U-MIDAS (DGP: HF-VMA(3); aggregation method: addition)

Parameters				Relative performance (K_{max}=3)						Relative performance (K_{max}=12)					
θ_l	θ_h	δ_l	δ_h	$Mean$	$10th$ $percentile$	$25th$ $percentile$	$Median$	$75th$ $percentile$	$90th$ $percentile$	$Mean$	$10th$ $percentile$	$25th$ $percentile$	$Median$	$75th$ $percentile$	$90th$ $percentile$
0	-0.5	0.1	0	0.997	0.946	0.972	0.994	1.020	1.056	0.996	0.942	0.971	0.994	1.020	1.056
0	-0.5	0.5	0	1.010	0.866	0.920	0.967	1.013	1.079	1.025	0.902	0.936	0.973	1.005	1.056
0	-0.5	1	0												
0	-0.7	0.1	0	0.997	0.941	0.967	0.995	1.024	1.058	0.995	0.933	0.967	0.995	1.022	1.061
0	-0.7	0.5	0	0.941	0.820	0.879	**0.943**	0.999	1.058	0.972	0.896	0.929	0.973	1.007	1.051
0	-0.7	1	0												
0	-0.9	0.1	0	0.992	0.941	0.965	0.991	1.017	1.048	0.991	0.940	0.966	0.990	1.017	1.049
0	-0.9	0.5	0	0.923	0.789	0.857	**0.936**	0.999	1.055	0.966	0.865	0.916	0.970	1.020	1.070
0	-0.9	1	0	0.749	0.565	0.645	**0.732**	0.838	0.935	0.994	0.787	0.853	0.962	1.088	1.235
-0.5	0	0.1	0												
-0.5	0	0.5	0	0.919	0.811	0.860	**0.915**	0.969	1.025	0.913	0.805	0.856	**0.910**	0.968	1.020
-0.5	0	1	0	0.935	0.854	0.898	**0.934**	0.975	1.015						
-0.7	0	0.1	0	0.953	0.870	0.907	**0.946**	0.998	1.041	0.952	0.870	0.905	**0.946**	0.997	1.038
-0.7	0	0.5	0	0.820	0.712	0.754	**0.810**	0.873	0.955	0.819	0.701	0.744	**0.815**	0.876	0.954
-0.7	0	1	0	0.786	0.661	0.714	**0.770**	0.845	0.926	0.814	0.682	0.734	**0.798**	0.876	0.962
-0.9	0	0.1	0												
-0.9	0	0.5	0	0.817	0.697	0.745	**0.809**	0.879	0.945	0.818	0.702	0.751	**0.811**	0.874	0.938
-0.9	0	1	0	0.854	0.611	0.682	**0.810**	0.992	1.166	0.892	0.645	0.724	**0.849**	1.023	1.198
-0.5	-0.5	0.1	0	0.961	0.861	0.915	0.957	1.004	1.057	0.959	0.860	0.912	0.955	1.003	1.053
-0.5	-0.5	0.5	0	0.865	0.743	0.803	**0.866**	0.925	0.978	0.981	0.799	0.838	**0.885**	0.943	0.997
-0.5	-0.5	1	0	0.741	0.603	0.655	**0.731**	0.812	0.897	0.900	0.795	0.843	**0.898**	0.954	1.007
-0.7	-0.7	0.1	0	0.888	0.777	0.828	**0.882**	0.943	1.013	0.887	0.775	0.826	**0.882**	0.942	1.009
-0.7	-0.7	0.5	0	0.956	0.877	0.908	**0.949**	0.997	1.041	0.954	0.873	0.908	**0.946**	0.995	1.039
-0.7	-0.7	1	0	0.609	0.467	0.528	**0.592**	0.668	0.766	0.863	0.705	0.764	**0.840**	0.919	1.005
-0.9	-0.9	0.1	0												
-0.9	-0.9	0.5	0	0.786	0.648	0.706	**0.780**	0.860	0.934	0.828	0.691	0.749	**0.823**	0.901	0.979
-0.9	-0.9	1	0	0.636	0.429	0.519	**0.635**	0.746	0.851	0.838	0.592	0.668	**0.829**	0.963	1.101

In bold rows where median MSFE of TF-MIDAS is at least 5% better than U-MIDAS.