

Nowcasting distributions: a functional MIDAS model

Massimiliano Marcellino [†] Andrea Renzetti [‡] Tommaso Tornese [§]

This Draft: November 11, 2024

Abstract

We propose a functional MIDAS model to leverage high-frequency information for forecasting and nowcasting distributions observed at a lower frequency. We approximate the low-frequency distribution using Functional Principal Component Analysis and consider a group lasso spike-and-slab prior to identify the relevant predictors in the finite-dimensional SUR-MIDAS approximation of the functional MIDAS model. In our application, we use the model to nowcast the U.S. households' income distribution. Our findings indicate that the model enhances forecast accuracy for the entire target distribution and for key features of the distribution that signal changes in inequality.

J.E.L Classification Code: C32; E32

Keywords: Nowcasting; Functional PCA; SUR; MIDAS; Inequality

*Andrea Renzetti and Massimiliano Marcellino thank MUR-Prin 2022 - Prot. 20227YZ9JK, financed by the European Union - Next Generation EU, for partial financial support. We thank Vasco Botelho, Michael Pfarrhofer, Anna Simoni and Jonas Striakus for helpful comments.

[†]Bocconi University and BAFFI-CAREFIN Center; email: massimiliano.marcellino@unibocconi.it

[‡]Bocconi University and BAFFI-CAREFIN Center; email: andrea.renzetti2@unibo.it.

[§]Università Cattolica del Sacro Cuore, Milano; email: tommaso.tornese@unicatt.it

1 Introduction

Economists, policymakers, and social scientists have long recognized the profound implications of income distribution for social stability, economic growth, and overall well-being. The ongoing shifts in income distribution are central to understanding the evolution of economic inequality. Additionally, the income distribution influences aggregate demand through consumption and investments, as well as the demand for financial assets (Aiyagari 1994; Kaplan et al. 2018; Bayer et al. 2019; Bilbiie et al. 2023). Therefore, closely monitoring the evolution of household income distribution is crucial, not only for tracking household income inequality per-se, but also for an accurate assessment of the current state of the economy. Real-time monitoring of income distribution provides critical insights that can guide effective policy interventions and promote equality, economic development and stability.

Most of the times, however, for what concerns the distribution of households' income we only have imperfect knowledge of the present state and even of the recent past. Comprehensive surveys on households' income are often released with considerable lag, since gathering data from a large and diverse sample of households is time-consuming. Surveys often require in-person interviews, phone calls, or mailed questionnaires, which can take several months to complete. This is especially true for large surveys that involve a representative sample of households and provide information on the annual stream of income, deriving not only from wages and salaries, but also from self employed income, social security benefits, interest and dividends, retirement income, unemployment compensations etc.

The importance of monitoring real-time changes in income distribution necessitates seeking timely signals from various indicators available before the release of official data. In this work we introduce a functional MIDAS (MIXed DATA Sampling) model for nowcasting the low frequency households' income distribution leveraging high frequency macroeconomic and financial indicators.¹ To exploit high-frequency macroeconomic indicators for nowcasting the households' income distribution, we face three main challenges. First, we need a finite dimensional approximation of the continuous distribution of household income. Second, we have to match the high frequency macroeconomic indicators with the low-frequency approximation of the distribution of income. Third, we might want to identify the relevant predictors from a potentially large number of macroeconomic and financial time series. These facts can easily cause a proliferation

1. For an extensive treatment of MIDAS models see Ghysels et al. (2025).

of parameters, potentially leading to overfitting with subsequent negative effects on forecast accuracy, particularly due to the small sample size resulting from annual or biannual observations in these types of household surveys. We address the first challenge using functional principal component analysis (Ramsay et al. 2005), which allows us to approximate the distribution of income with a relatively small number of basis functions. Approximating the distribution of income using functional principal components leads to a seemingly unrelated MIDAS regression (SUR-MIDAS) approximation of the functional MIDAS model. Still, when considering a possibly large number of high frequency indicators and a meaningful lag dynamics for these indicators, the SUR-MIDAS model is highly-parameterized. For this purpose, we extend the MIDAS group lasso spike and slab adaptive prior in Mogliani et al. (2021) to a SUR framework. Since this prior is constructed to set exactly to zero the coefficients of the irrelevant predictors, this approach allows us to explore the usefulness of a potentially large number of predictors while performing variable selection. For the specification of the prior, we consider a re-parameterization of the original SUR model which enables an efficient equation by equation estimation of the parameters of the model. Thus, the paper also contributes to the recent MIDAS literature by extending the sparse-group LASSO regression-based approach (Babii et al. 2021; Mogliani et al. 2021; Babii et al. 2024) to a SUR framework.

Our model is a direct forecasting tool that matches high-frequency information from macroeconomic and financial indicators to predict the entire household income distribution before it is officially released. In general, although it does not enable predictions for the micro-level distribution at a high frequency, the MIDAS framework allows leveraging high-frequency information for nowcasting the low-frequency micro-level income distribution. This applies both when the micro variable reports income for the fourth quarter of the current year and when they refer to the entire year, as in the application we consider. Indeed, since the MIDAS approach does not require filtering out missing observations of the micro-variable’s low-frequency distribution, it avoids the issue of aggregating latent high-frequency distributions, as it would be required in the mixed frequency framework of Schorfheide et al. (2015). This circumvents the need to impose additional assumptions when approximating the cross-sectional distributions.

In the application, we use the model for nowcasting the distribution of households’ income in the United States from the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS). This annual survey on yearly household income is released in March of the following year. This data provides a comprehensive measure of the financial resources

accrued by a household in a given year, and it is usually used for analyzing economic well-being and poverty status among different demographic groups in the United States. To investigate the usefulness of high frequency macroeconomic indicators for nowcasting the distribution of household income, we set up a pseudo-real-time forecasting exercise. The set of predictors comprehends quarterly indicators concerning the evolution of aggregate income, employment, financial accounts, government expenditure, consumption, interest rates and stock prices within the year. In the application, we find that leveraging quarterly macroeconomic and financial indicators with our model enhances forecast accuracy about the yearly cross-sectional distribution of household income for the US. In particular, we find that exploiting high frequency information allows to enhance forecast accuracy of key features of the distribution that signal changes in inequality.

The rest of the article is organized as follows. In Section 2 we introduce the functional MIDAS model and the group lasso-type prior for the parameters of its finite dimensional SUR-MIDAS approximation. To assess the small-sample performance of our modelling strategy and the prior specification, in section 3 we set-up a Monte Carlo simulation, which highlights the merits of our approach in realistic settings. Section 4 presents the application of the SUR-MIDAS to the problem of nowcasting the cross-sectional distribution of household income in the US released in the March ASEC of the CPS. Section 5 summarizes the main findings and concludes. The appendix presents additional theoretical and empirical results.

2 A Functional-MIDAS model for nowcasting distributions

We analyze the dynamics of the cross-sectional distributions of income by modelling the associated Log-Quantile Density (LQD) function, i.e the first derivative of the quantile function corresponding to the distribution of interest (Parzen 1979; Jones 1992). That is:

$$q_t(\tau) = \log \left(\frac{\partial Q_t(\tau)}{\partial \tau} \right) \quad (1)$$

where $Q_t(\tau)$ is the quantile function mapping $\tau \in [0, 1] \rightarrow [\underline{x}, \bar{x}]$ at each time t , for $t = 1, \dots, T$. As discussed by Petersen et al. 2016 and Huber et al. 2024, the use of the LQD function is convenient since $q_t(\tau)$ does not need to integrate to one or to obey non-negativity constraints, as it would be the case for the probability density function, or satisfy monotonicity constraints,

which would be the case for the cumulated density function or the quantile function. As enforcing such constraints in a dynamic linear model is extremely challenging, modelling the LQD simplifies considerably the econometric analysis of the time variation in the distribution of interest. The explicit mapping from the probability density function $f(x)$ into the LQD function, is defined by $\psi_Q(f)(\tau) = -\log(f(Q(\tau)))$, where $Q(\tau)$ is the quantile function. To model the dynamic interaction between the low frequency cross-sectional distribution of income and the high frequency macroeconomic indicators, we specify the following functional MIDAS regression model:

$$q_t(\tau) = c_q(\tau) + \sum_{l=0}^{p_x-1} \mathbf{B}_{qx,l}(\tau)L(l/m)\mathbf{x}_t^{(m)} + \sum_{l=1}^{p_q} \int_0^1 B_{qq,l}(\tau,\tau')L(l)q_t(\tau')d\tau' + u_{q,t}(\tau) \quad (2)$$

where $\mathbf{x}_t^{(m)}$ is the $n_x \times 1$ vector of high-frequency macroeconomic indicators, $L(\cdot)$ is the lag operator, such that $L(1/m)\mathbf{x}_t = \mathbf{x}_{t-1/m}^{(m)}$ and $L(1)q_t(\tau) = q_{t-1}(\tau)$. The function $q(\tau)$ and the high-frequency macroeconomic indicators $\mathbf{x}_t^{(m)}$ are assumed to be sampled at different frequencies. For instance, when the cross-sectional distribution is observed yearly and the macroeconomic indicators are observed quarterly, then $m = 4$.² Hence, in the model, the LQD function is expressed as a combination of p_x lags of the high frequency macroeconomic indicators and its own p_q lags. We assume that the LQD function admits the Karhunen-Loève expansion

$$q_t(\tau) = \mu(\tau) + \sum_{k=1}^{\infty} \sqrt{\lambda_k} z_{t,k} h_k(\tau) \quad (3)$$

and we approximate $q_t(\tau)$ by truncating the infinite sum in (3) at a level K , namely:

$$q_t(\tau) = \mu(\tau) + \sum_{i=1}^K h_i(\tau) f_{t,i} = \mu(\tau) + \mathbf{h}_K(\tau)' \mathbf{f}_{t;K} \quad (4)$$

where $\mathbf{h}_K(\tau)$ is a $K \times 1$ vector of basis functions, while $\mathbf{f}_{t;K}$ is a $K \times 1$ vector of coefficients or scores associated to the basis functions. Once the LQD function is approximated using K basis functions, the functional MIDAS model in (2) can be rewritten as a SUR-MIDAS (Seemingly Unrelated Regression - MIDAS) model for $\mathbf{f}_{t;K}$, that is

$$\mathbf{f}_{t;K} = \phi_0 + \sum_{l=0}^{p_x-1} \Phi_{fx,l} L(l/m)\mathbf{x}_t^{(m)} + \sum_{l=1}^{p_q} \Phi_{ff,l} L(l)\mathbf{f}_{t;K} + \mathbf{u}_{f,t} \quad (5)$$

2. Note that we are considering the same number of lags for all the high frequency macroeconomic indicators n_x . This is for ease of exposition.

with $E[\mathbf{u}_{f,t}] = \mathbf{0}$ and $E[\mathbf{u}_{f,t}\mathbf{u}_{f,t}'] = \mathbf{\Omega}$.³ It is easy to see that when the number of basis functions in the approximation of the log-quantile-density function K is big, the SUR-MIDAS model becomes highly-parameterized. To approximate the LQD functions $q_t(\tau)$ using a relatively small number of basis functions, we resort to Functional Principal Component Analysis (FPCA, Ramsay et al. 2005). While this approach allows to obtain a SUR-MIDAS model for a vector $\mathbf{f}_{t;K}$ of moderate dimensions, in our application we still face two challenges. First, we might consider incorporating a wide array of macroeconomic indicators and then select which of them are most relevant. Second, we might need a potentially large number of lags of the high frequency macroeconomic indicators to capture a meaningful lag dynamics. For example, in a typical application in which the cross-sectional distribution is observed at the yearly frequency and the macroeconomic indicators are observed at the quarterly frequency, including an extra macroeconomic indicator with the corresponding observations for just one year leads to $4K$ additional coefficients in the model. In general, since the model features Kp_xn_x coefficients on the high frequency macroeconomic indicators and K^2p_q coefficients on the lags of $\mathbf{f}_{t;K}$, the number of parameters increases as the number of high frequency indicators n_x or the number of their lags p_x increases, but also when the number of lags of the log-quantile-density function p_q increases. This proliferation of parameters can easily lead to overfitting since working with yearly cross-sectional income distributions typically implies small sample sizes in most applications. We address this issue by considering a group lasso spike and slab type prior for the coefficients in the MIDAS-SUR model (Xu et al. 2015; Mogliani et al. 2021). Section 2.1 below, presents the details concerning the approximation of the LQD function with FPCA, while section 2.2 presents the group lasso-type prior for the SUR-MIDAS model; section 2.3 discusses forecasting and nowcasting, and section 2.4 introduces the estimation algorithm.

2.1 Approximation of the LQD function by FPCA

Our estimation strategy follows a two-step approach. First, we approximate the low-frequency distributions using functional principal component analysis. Next, we estimate the SUR MIDAS model to generate nowcasts of the factors, and consequently, of the corresponding distributions.⁴ In this section, we detail the approximation of the targeted micro-variable distribu-

3. Appendix A.1 reports the steps involved to go from the Functional MIDAS model to its SUR-MIDAS model representation.

4. Conditioning on the estimates of the eigenfunctions from FPCA in the first step, the factors can also be treated as random and estimated jointly with the parameters of the SUR MIDAS model in the Gibbs Sampler by

tion through FPCA. The basis function approximation in equation (4) allows to approximate non-parametrically potentially very flexible distributions. FPCA can be used to determine the “optimal basis functions” for approximating functional data. These basis functions are derived to capture the maximum variance in the data with a given number of components, ensuring that the representation is both efficient and comprehensive. In the context of approximating distribution functions, Petersen et al. (2016) has pointed out the advantages of exploiting the LQD transformation. The LQD transformation ensures that the original density functions are mapped into a linear Hilbert space, which is essential for the proper application of FPCA. This approach retains the interpretability of the principal components while overcoming the challenges posed by the nonlinear constraints intrinsic in the space of density functions. Additionally, the LQD transformation leads to more efficient and meaningful representations of the variability in a sample of densities, as it captures both vertical and horizontal variations more effectively than direct FPCA on densities. Observations from the LQD functions $q_t(\tau)$ are obtained by evaluating

$$q(\tau) = -\log(\hat{f}(Q(\tau))) \quad (6)$$

on a set of grid points $\tau_1, \dots, \tau_{N^{grid}}$. In Equation (6), $\hat{f}(\cdot)$ is a kernel smoothed density estimate of the probability density function namely:

$$\hat{f}(x) = \frac{1}{N^{grid}h} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right) \quad (7)$$

We resort to FPCA in order to approximate $q_t(\tau)$ in equation (4). Specifically, the basis functions $\mathbf{h}_K(\boldsymbol{\tau})$ and the scores $\mathbf{f}_{t;K}$ are obtained by static FPCA of the deviation of $q_t(\tau)$, observed over $t = 1, \dots, T$, from their mean. More precisely, to represent the continuous functions $q_t(\tau)$ in a finite dimensional data matrix $\mathbf{Q} \in \mathbb{R}^{N^{grid} \times T}$, we center the data matrix by subtracting from each LQD function $q_t(\tau)$ evaluated on a grid $\tau = \tau_1, \dots, \tau_{N^{grid}}$ a simple estimate of their mean obtained as $\hat{\mu}(\tau) = \frac{1}{T} \sum_{t=1}^T q_t(\tau)$, to get \mathbf{Q} . Then we apply Singular Value Decomposition (SVD) to the centered matrix, namely

$$\mathbf{Q} = \mathbf{U}\mathbf{S}\mathbf{V}^T \quad (8)$$

where \mathbf{U} and \mathbf{V} are orthogonal matrices, and \mathbf{S} is a diagonal matrix containing the singular values (we omit the dependence on τ just for notation convenience). To obtain the basis functions

the Kalman filter (Carter et al. 1994) or by the precision sampler (Chan et al. 2009).

$\mathbf{h}(\boldsymbol{\tau})$, we select K columns from \mathbf{V} , which represent the eigenbasis associated with the K largest eigenvalues. The $\mathbf{f}_{t;K}$ in (4) are therefore given by the principal component scores corresponding to the selected K eigenbasis. Petersen et al. (2016) demonstrates the consistency of these estimates and we refer to this paper for rates of convergence.

2.2 A Group Lasso Shrinkage prior for the SUR-MIDAS model

Once the LQD functions are approximated by FPCA, the functional MIDAS model in (2) can be rewritten as the SUR-MIDAS model for the principal component factors $\mathbf{f}_{t;K}$ in (5). As anticipated above, this model is highly parameterized. In this section we define a group lasso-type shrinkage prior for the coefficients in the SUR-MIDAS model. This prior is meant to select the most relevant predictors and, at the same time, shrinking the coefficients within the selected groups to zero. More in detail, this prior extends the MIDAS adaptive group lasso prior of Mogliani et al. (2021) to a SUR framework. We collect the lags of the high frequency macroeconomic indicators and the lags of the low frequency factors in the vector $\mathbf{x}_t = [\text{vec}([\mathbf{x}_{t-1/m}^{(m)} \dots \mathbf{x}_{t-p_x/m}^{(m)}]'), \text{vec}([\mathbf{f}_{t-1}, \dots, \mathbf{f}_{t-p_q}]')']'$. Then we rewrite the model (5) as

$$\mathbf{f}_t = \boldsymbol{\Phi}' \mathbf{x}_t + \mathbf{u}_t \quad \mathbf{u}_t \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \quad (9)$$

where we are setting $h = 0$ (nowcasting) and consider $\phi_0 = 0$ just for the sake of the exposition.⁵ In this notation, the matrix of coefficients $\boldsymbol{\Phi}$ of dimension $(n_x p_x + K p_q) \times K$ is storing the coefficients on both the lags of the high frequency indicators and the low frequency factors in all the K equations of the SUR model. In equation (9) we assume normality of the error term \mathbf{u}_t and allow for possibly correlated errors across equations, meaning that $\mathbb{E}[u_{1t}^2] = \omega_i^2$ and $\mathbb{E}[u_i u_k] = \omega_{ik}$ for $k \neq i$. Rewriting the SUR model equation by equation we have

$$\begin{aligned} f_{1t} &= \mathbf{x}_t' \boldsymbol{\phi}_1 + u_{1t} \\ f_{2t} &= \mathbf{x}_t' \boldsymbol{\phi}_2 + u_{2t} \\ &\vdots \\ f_{Kt} &= \mathbf{x}_t' \boldsymbol{\phi}_K + u_{Kt} \end{aligned} \quad (10)$$

5. For ease of exposition we also suppress the K subscript in $\mathbf{f}_{t;K}$ and just write it as \mathbf{f}_t

where the vector ϕ_i for $i = 1, \dots, K$ is the i^{th} column of Φ . We re-parameterize the SUR model following Zellner et al. (2010) as follows

$$\begin{aligned}
f_{1t} &= \mathbf{x}_t' \phi_1 + \varepsilon_{1t} & \varepsilon_{1t} &= u_{1t} \sim \mathcal{N}(0, \sigma_1^2) \\
f_{2t} &= \mathbf{x}_t' \phi_2 + \alpha_{21} u_{1t} + \varepsilon_{2t} & \varepsilon_{2t} &\sim \mathcal{N}(0, \sigma_2^2) \\
&\vdots & & \\
f_{Kt} &= \mathbf{x}_t' \phi_K + \alpha_{K,1} u_{1t} + \dots + \alpha_{K,K-1} u_{K-1t} + \varepsilon_{Kt} & \varepsilon_{Kt} &\sim \mathcal{N}(0, \sigma_K^2)
\end{aligned} \tag{11}$$

This re-parameterization is convenient since it allows to rewrite the system of equations of the SUR model as a set of univariate regressions, with uncorrelated error terms ε_{it} . Clearly, the correlation across the equations remains evident, originating from the product of the $u_{i,t}$ terms and the coefficients $\alpha_{i,c}$ for $i = 2, \dots, K$ and $c = 1, \dots, i-1$. We specify a prior for the elements of $\alpha_{i,\cdot} = [\alpha_{i,1}, \dots, \alpha_{i,i-1}]$ and for the variances $\sigma_1^2, \dots, \sigma_K^2$ which implies a symmetric prior for the variance covariance matrix of the error term in the original model Ω . More specifically, we assume

$$(\alpha_{i,c} | \sigma_i^2) \sim \mathcal{N}\left(0, \frac{\sigma_i^2}{s_c^2}\right) \quad 1 \leq c \leq i = 2, \dots, K \tag{12}$$

$$\sigma_i^2 \sim \mathcal{IG}\left(\frac{v_0 + i - K}{2}, \frac{s_i^2}{2}\right) \quad i = 1, \dots, K \tag{13}$$

It can be shown that this specification implies an *Inverse-Wishart* prior for the variance covariance matrix of the error term in the original model (5), namely⁶

$$\Omega \equiv \mathbb{E}[\mathbf{u}_t \mathbf{u}_t'] = \tilde{\mathbf{A}}^{-1} \Sigma (\tilde{\mathbf{A}}^{-1})' \sim \mathcal{IW}(\mathbf{S}, v_0) \tag{14}$$

where $\Sigma = \text{diag}(\sigma_1^2, \dots, \sigma_K^2)$, $\tilde{\mathbf{A}}$ is the lower triangular matrix with unit elements on the main diagonal and rows given by $\tilde{\alpha}_{i,\cdot} = [-\alpha_{i,1}, \dots, -\alpha_{i,i-1}]$ and $\mathbf{S} = \text{diag}(s_1^2, \dots, s_K^2)$. Hence, this specification keeps the symmetry of the prior distribution for the variance-covariance matrix of the error term in the original model. At the same time, the specification allows to directly specify a prior for the coefficients in Φ . In this respect, our SUR-MIDAS model features an unrestricted lag-dynamics for both the high frequency macroeconomic indicators and for the low frequency factors. To specify a group lasso-type prior for the elements in ϕ_i i.e. the coefficients in the $i = 1, \dots, K$ equations of the SUR-MIDAS model, the natural approach would be to define n_x groups,

6. The proof of this result can be found in Chan (2022).

one for each high frequency macroeconomic indicator, and K groups, one for each low frequency factor. This makes a total of $G = n_x + K$ groups with n_x groups of dimension p_x (the number of lags for each high frequency regressor) and K groups of dimension p_q (the number of lags for each low frequency factor). Alternatively, as in Mogliani et al. (2021), one can reduce the number of unknown parameters by restricting the lag-dynamics of the high frequency macroeconomic indicators $\sum_{l=0}^{p_x-1} L(l/m)\mathbf{x}_t^{(m)}$ and low-frequency factors $\sum_{l=1}^{p_q} L(l)\mathbf{f}_{t;K}$, exploiting the Almon lag polynomial. More specifically, the direct representation of the Almon lag polynomial allows to keep the model linear in the parameters, by just considering a transformation of the regressors \mathbf{x}_t . In particular, considering an Almon lag polynomial leads to the following transformation of the regressors

$$\mathbf{z}_{j,t} = \mathbf{W}\mathbf{x}_{j,t} \quad (15)$$

where $\mathbf{x}_{j,t} = [x_{jt-1/m}^{(m)}, \dots, x_{j,t-p_x/m}^{(m)}]'$ for the lags of the high frequency variables and $\mathbf{x}_{j,t} = [x_{jt-1}, \dots, x_{j,t-p_q}]'$ for the lags of the low frequency factors and \mathbf{W} is the polynomial weighting matrix of this transformation. This approach allows to further reduce the dimension of the groups made of the lags of the high frequency indicators and of the lags of the low frequency factors to $p_a + 1 - r_a$ where p_a is the order of the Almon lag polynomial and r_a are the number of restrictions on the shape of the Almon lag polynomial.⁷ Exploiting the direct representation of the Almon lag polynomial, we rewrite the SUR-MIDAS model as:

$$\begin{aligned} f_{1t} &= \mathbf{z}'_t\boldsymbol{\theta}_1 + \varepsilon_{1t} & \varepsilon_{1t} &\sim \mathcal{N}(0, \sigma_1^2) \\ f_{2t} &= \mathbf{z}'_t\boldsymbol{\theta}_2 + \alpha_{21}u_{1t} + \varepsilon_{2t} & \varepsilon_{2t} &\sim \mathcal{N}(0, \sigma_2^2) \\ &\vdots & & \\ f_{Kt} &= \mathbf{z}'_t\boldsymbol{\theta}_K + \alpha_{K,1}u_{1t} + \dots + \alpha_{K,K-1}u_{K-1t} + \varepsilon_{Kt} & \varepsilon_{Kt} &\sim \mathcal{N}(0, \sigma_K^2) \end{aligned} \quad (16)$$

where the vector $\boldsymbol{\theta}_i = [\boldsymbol{\theta}_{i,1}', \dots, \boldsymbol{\theta}_{i,G}']'$ collects all the parameters of the Almon lag polynomial of each of the $G = (n_x + K)$ groups in the i^{th} equation of the SUR model. The vector \mathbf{z}_t is instead storing the transformed regressors. Note that each vector $\boldsymbol{\theta}_{ij}$ is of dimension $(p_a + 1 - r_a) \times 1$ as, for ease of exposition, we are considering the same Almon lag polynomial order and the same

7. The weighting matrix \mathbf{Q} is of dimension $(p_a + 1 - r_a) \times p_x$ for the transformation of the high frequency variables and of dimension $(p_a + 1 - r_a) \times p_q$ for the transformation of the low frequency factors. Note that just for ease of the exposition we are considering the same degree of the Almon lag polynomial for all the groups. Clearly, it is also possible to keep an unrestricted lag dynamics on the lags of the low-frequency factors by just considering a linear transformation where $\mathbf{Q} = \mathbf{I}_{p_q}$ such that the dimension of the groups of each low frequency factor remains p_q .

number of restrictions for all the groups. Once we have obtained the groups and defined $\boldsymbol{\theta}_{ij}$ as the vector of coefficients in the i^{th} equation for the j^{th} group we specify the following prior (Xu et al. 2015; Mogliani et al. 2021)

$$p(\boldsymbol{\theta}_{ij} | \sigma_i^2, \tau_{ij}^2, \pi_{i0}) = (1 - \pi_{i0})\mathcal{N}(\mathbf{0}, \sigma^2 \tau_{ij}^2 \mathbf{I}_{g_{ij}}) + \pi_{i0} \delta_0(\boldsymbol{\theta}_{ij}) \quad (17)$$

$$p(\tau_{ij}^2) \sim \text{Gamma}\left(\frac{g_{ij} + 1}{2}, \frac{\lambda_{ij}^2}{2}\right) \quad (18)$$

$$\pi_{i0} \sim \text{Beta}(c, d) \quad (19)$$

for $i = 1, \dots, K$ and $j = 1, \dots, G$. This prior introduces two types of shrinkage effects: a point mass at zero, resulting in exact zero coefficients (the spike component), and a Group Lasso prior applied to the slab component. Therefore this prior enhances variable selection at the group level while simultaneously shrinking the coefficients within the selected groups.

2.3 Forecasting

The MIDAS regression is a direct forecasting tool which results in different forecasting models for each forecast horizon (Ghysels et al. 2004). Collecting the parameters of the Almon lag polynomial for all the groups in the i^{th} equation of the SUR-MIDAS model in $\boldsymbol{\theta}_i = [\boldsymbol{\theta}_{i,1}', \dots, \boldsymbol{\theta}_{i,G}']'$ and defining $\boldsymbol{\Theta}$ as the matrix stacking the vectors $\boldsymbol{\theta}_i$ for $i = 1, \dots, K$ on its columns, we can write the h -step ahead direct forecast of the low frequency factors as

$$\mathbf{f}_{t+h} = \boldsymbol{\Theta}' \mathbf{z}_t + \mathbf{u}_{t+K} \quad \mathbf{u}_{t+K} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}) \quad (20)$$

The h -step ahead direct forecast of the LQD function is therefore

$$q_{t+h}(\tau) = \mu(\tau) + \mathbf{h}(\boldsymbol{\tau})' \mathbf{f}_{t+h}. \quad (21)$$

From the forecast of the LQD function we obtain the corresponding distribution function, as in Kokoszka et al. (2019). More in detail, the backward mapping recovers the h -step ahead predicted density function from the log-quantile density function through $f(x) = \exp\{-\psi_Q(f)(F(x))\}$. In this regard, our MIDAS approach allows to recover the distribution function regardless of whether the micro-variable is a flow or a stock. Unlike the standard mixed-frequency literature (Schorfheide et al. 2015), we do not treat the high-frequency distribution of the micro-level

variable as latent. Therefore, we avoid the issue of aggregating missing high-frequency distributions, which can be non-trivial task, as noted by Chang et al. (2022). For instance, in the case of flow variables, our method can accommodate both low-frequency observations relative to the last high-frequency period and those spanning the entire low-frequency period. In other words, when predicting the household-level income distribution observed at a yearly frequency using quarterly predictors, our approach works regardless of whether the micro variable reports income for the fourth quarter of the current year or for the entire year. In the latter case, as with the income variable from the CPS ASEC in our application, specifying an aggregation rule for the missing quarterly distribution within the mixed-frequency framework of Schorfheide et al. 2015 would be complex and might require a more structured approach. This complexity arises because the observed yearly income distribution is a convolution of four unobserved quarterly income distributions and would require imposing additional structure by means of an appropriate aggregation rule.⁸ Hence, although it does not enable high-frequency predictions for the micro-level distribution, the MIDAS framework allows to impose less structure when approximating low frequency distributions and leverage high frequency information.

2.4 Bayesian Inference

As anticipated above, inference proceeds in two steps. First we approximate the LQD function using FPCA and obtain an estimate of the factors \mathbf{f}_t . Then we estimate the SUR-MIDAS model using a Gibbs sampler that iteratively samples from the conditional posterior distributions of the parameters in model (16), namely:

1. Sample from $p(\boldsymbol{\theta}_{ij}|\cdot)$ for $i = 1, \dots, K$ and $j = 1, \dots, G$.
2. Sample from $p(\boldsymbol{\alpha}_{i,\cdot}|\cdot)$ for $i = 2, \dots, K$
3. Sample from $p(\sigma_i^2|\cdot)$ for $i = 1, \dots, K$
4. Sample from $p(\tau_{ij}|\cdot)$ for $i = 1, \dots, K$ and $j = 1, \dots, G$
5. Sample $p(\pi_{i0}|\cdot)$ for $i = 1, \dots, K$

As anticipated above, the specification of the priors on the re-parameterized version of the model (11) enables the use of an efficient posterior sampling algorithm to produce inference

⁸ De Polis et al. (2024) recently consider a mixed frequency functional VAR, treating the high frequency cross-sectional distributions as missing and imposing such intertemporal restrictions in the aggregation rule.

about coefficients of the model equation by equation. In particular, to sample the coefficients in Step 1 we exploit the triangular algorithm proposed in Carriero et al. (2019) and corrected in Carriero et al. (2022). The formulas of the conditional posterior distributions are reported in the appendix A.2. To set the hyper-parameter λ_{ij}^2 , we follow the approach in Mogliani et al. (2021) which is an Empirical Bayes approach that relies on the so-called internal adaptive MCMC algorithms in Atchadé et al. (2011).⁹

3 Monte Carlo simulation

We perform a Monte Carlo simulation to investigate the gains of exploiting high frequency variables for nowcasting a distribution observed at a lower frequency with our approach. In particular, we assume that the true LQD function of the low frequency distribution we aim to forecast has the following representation

$$q_t(\tau) = \mu(\tau) + \sum_{i=1}^3 h_i(\tau) f_{t,i} = \mu(\tau) + \mathbf{h}(\tau)' \mathbf{f}_t \quad (22)$$

We assume that the probability density function associated to this LQD function is defined on the support $(0, 10)$. We consider the following SUR MIDAS dynamics for the elements of the vector of factors \mathbf{f}_t

$$f_{it} = \alpha_i + \sum_{j=1}^{30} \beta_{ij} \sum_{c=0}^{24-1} \tilde{B}(c; \theta) L_c^{(m)} x_{j,t-h}^{(m)} + \sum_{l=1}^2 \phi_{il}' \mathbf{f}_{t-l} + u_{i,t} \quad (23)$$

$$x_{j,t}^{(m)} = \mu + \rho x_{j,t-1/m}^{(m)} + \epsilon_{j,t} \quad (24)$$

where $i = 1, \dots, 3$ is the number of factors and therefore of equations in the SUR-MIDAS representation of the functional MIDAS model. $j = 1, \dots, 30$ is the number of high frequency variables while 24 is the lag length of the high frequency variables and 2 is the number of lags for the low frequency factors. The high frequency variables are sampled at frequency $m = 4$, to mimic yearly low frequency observations and quarterly high frequency observations as in our application. The lag-polynomial $\tilde{B}(c; \theta)$ denotes the normalized weights (i.e., summing up to 1). To set the weighting scheme we consider the DGP 1 in Mogliani et al. (2021). The lag dynamics

9. Alternatively we also report the formula of the conditional posterior of λ_{ij}^2 when treating it as random variable with Gamma prior distributions

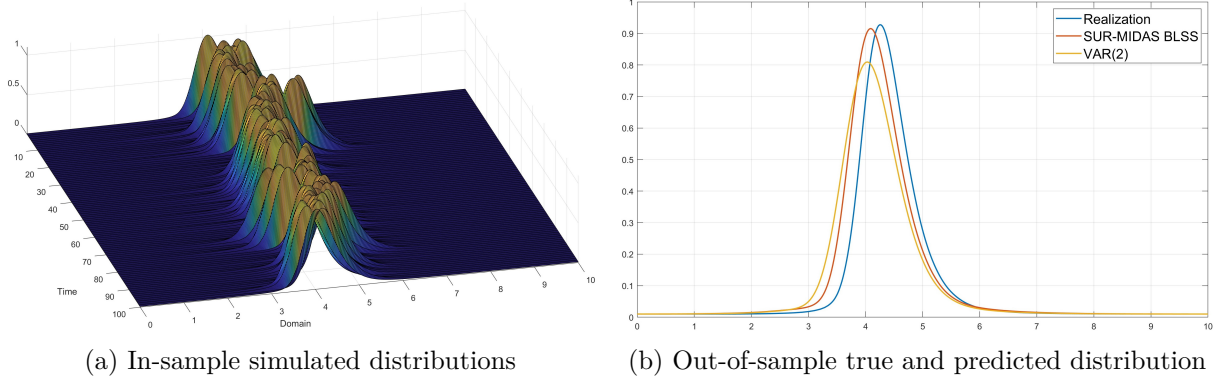
of the low frequency factors is instead left unrestricted. The error terms $\mathbf{u}_t := (u_{1,t}, \dots, u_{K,t})'$ and $\boldsymbol{\epsilon}_t := (\epsilon_{1,t}, \dots, \epsilon_{n_x,t})'$ are i.i.d. with distribution

$$\begin{pmatrix} \mathbf{u}_t \\ \boldsymbol{\epsilon}_t \end{pmatrix} \sim \mathcal{N} \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \boldsymbol{\Omega} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_\epsilon \end{bmatrix} \right), \quad (25)$$

Details on the other parameters of the DGP are in the Appendix A.3. The basis functions in (22) are selected as the eigenbasis corresponding to the three largest eigenvalues obtained through FPCA on the LQD function of a collection of *Skew-t* distributions (Azzalini et al. 2003). Each distribution is characterized by different location, shape, scale, tail thickness and it is truncated on the support (0,10).

As we want to assess whether our approach allows to correctly select the relevant high frequency predictors among the thirty high frequency indicators, we assume a sparse structure for the β_{ij} in (23). In other words, in each equation of the SUR-MIDAS model, only a subset of predictors is relevant for predicting the low-frequency factors and therefore changes in the distribution over time. We consider two sample sizes $T = \{60, 120\}$ and the number of the Monte Carlo simulations is 500. Figure 1, panel a) shows the time series of the simulated distributions in one replication of the Monte Carlo. Figure 1 panel b) shows the first out-of-sample realization of the true distribution, with the nowcast obtained from our functional VAR(2) model and a functional MIDAS model with LASSO type prior discussed in the text and a ridge type prior discussed in the appendix A.4. The point forecast of the distribution function is obtained by averaging across the posterior distribution draws. For ease of analysis we assume $h = 0$, i.e. a nowcasting model with high-frequency information fully matching the low frequency. While the SUR-MIDAS models leverage information from both the high frequency variables and the lags of the low frequency factors, the VAR(2) is just exploiting the information from the lags of the low frequency factors.

Figure 1: Output from one replication of the Monte Carlo



To assess the accuracy of the nowcasts across the Monte Carlo replications, we compare the Kullback-Leibler (KL) distance and the Hellinger Distance (HD) to the true probability density function, the quantile scores (QS) and the root mean squared error (RMSE) of some selected moments of the distribution such as mean, variance, skewness and kurtosis. As a measure of dispersion we consider the interquartile range while as a measure of inequality we consider the Gini and the coefficient of variation (CV).

Table 1 reports the results from the Monte Carlo simulation. Even in small sample size $T = 60$, the spike and slab prior correctly selects the relevant predictors in the model, outperforming both the SUR-MIDAS with ridge type prior and the VAR(2). The Monte Carlo simulation shows that exploiting high frequency information allows to improve the forecasts from the entire distribution and of key features that signal changes in inequality such as the Gini Index and the coefficient of variation.

Table 1: Forecast accuracy in the Monte Carlo Simulation

T=60							
	Avg KL	Avg HD	Avg QS5	Avg QS20	Avg QS50	Avg QS80	Avg QS95
VAR(2)	0.1651	0.1817	0.0760	0.0992	0.1195	0.1221	0.5795
RIDGE SUR-MIDAS	0.1204	0.1439	0.0713	0.0826	0.0934	0.1038	0.4422
BLASSO SUR-MIDAS	0.1029	0.1343	0.0645	0.0757	0.0861	0.0935	0.4188
	RMSE Mean	RMSE Variance	RMSE Skewness	RMSE Kurtosis	RMSE IR	RMSE GINI	RMSE CV
VAR(2)	0.1794	2.1970e-1	0.4830	2.4652	0.1129	0.0623	2.1684
RIDGE SUR-MIDAS	0.1648	1.6270e-1	0.3840	1.8103	0.0438	0.0437	1.7081
BLASSO SUR-MIDAS	0.1570	1.5640e-1	0.3908	1.8028	0.0492	0.0436	1.5925
T=120							
	Avg KL	Avg HD	Avg QS5	Avg QS20	Avg QS50	Avg QS80	Avg QS95
VAR(2)	0.1685	0.1794	0.0781	0.0995	0.1193	0.1247	0.5755
RIDGE SUR-MIDAS	0.1137	0.1393	0.0792	0.0846	0.0904	0.0975	0.4035
BLASSO SUR-MIDAS	0.0886	0.1208	0.0725	0.0740	0.0777	0.0838	0.3476
	RMSE Mean	RMSE Variance	RMSE Skewness	RMSE Kurtosis	RMSE IR	RMSE GINI	RMSE CV
VAR(2)	0.1795	2.1950e-1	0.4864	2.4725	0.0999	0.0631	2.1946
RIDGE SUR-MIDAS	0.1633	1.5130e-1	0.3969	1.7699	0.0414	0.0432	1.6184
BLASSO SUR-MIDAS	0.1383	1.3680e-1	0.3453	1.6040	0.0319	0.0384	1.4008

Notes: The table reports the results from Monte Carlo simulations with $T = 60$ in the first two sections and $T = 120$ in the last two sections. All values are rounded to four decimal places. Very small and very large numbers are written using scientific notation for readability.

4 Nowcasting household income distribution in the US

4.1 Data and design of the nowcasting exercise

We now use our functional MIDAS model to nowcast the yearly cross-sectional distribution of households' income in the United States. We focus on the variable `hhincome` from the Annual Social and Economic Supplement (ASEC) of the Current Population Survey (CPS). This variable aggregates the income received by all members of a household from various sources over the past year, including wages and salaries, self-employment earnings, interest, dividends, rents, and other forms of income like social security benefits, retirement pensions, and public assistance. In general, this variable provides a comprehensive measure of the financial resources available to a household, and it is typically used for analyzing economic well-being and poverty status among different demographic groups in the US. This annual survey on yearly household income is released in March of the following year. Throughout the year, up until the release date, information about the latest developments in both the real and the financial markets gradually becomes available. This data includes macroeconomic and financial indicators, which might be especially valuable for producing a more precise forecast of the yearly income distribution before

its official announcement in March. To leverage this information effectively, we include these high frequency indicators in a functional MIDAS model for the distribution of income. The set of quarterly macroeconomic indicators and their transformation can be found in table 2. The household-level nominal income given by the variable `hhincome` is normalized by dividing by 2/3 of nominal GDP per-capita and then transformed by computing the inverse hyperbolic sine function as in (Chang et al. 2021).¹⁰

Table 2: List of quarterly indicators

Variable	Mnemonic	Transformation
Real Gross Domestic Product	GDPC1	%Δ
Real Government Receipts	FGRECPtx	%Δ
Federal Government: Current Expenditures	FGEXPND	%Δ
Real Personal Consumption Expenditures	PCECC96	%Δ
Real Disposable Personal Income	DPIC96	%Δ
Real Exports of Goods	EXPGSC1	%Δ
Real Imports of Goods & Services	IMPGSC1	%Δ
All Employees: Total nonfarm	PAYEMS	%Δ
Civilian Labor Force Participation Rate	CIVPART	%Δ
Civilian Unemployment Rate	UNRATE	%Δ
Number of Civilians Unemployed for 27 Weeks and Over	UEMP27OV	%Δ
Average Weekly Hours Of Production And Nonsupervisory Employees	AWHNONAG	%Δ
Average (Mean) Duration of Unemployment	UEMPMEAN	Δ
Housing Starts	HOUST	%Δ
Real Average Hourly Earnings of Production and Nonsupervisory Employees: Total	AHETPIx	%Δ
Real Average Hourly Earnings of Production and Nonsupervisory Employees Construction	CES2000000008x	%Δ
Real Average Hourly Earnings of Production and Nonsupervisory Employees Manufacturing	CES3000000008x	%Δ
Consumer Price Index all items	CPIAUCSL	%Δ
S&P 500	SP_500	%Δ
Effective Federal Funds Rate	FEDFUNDS	Δ
10-Year Treasury Constant Maturity Rate (Percent)	GS10	Δ
Cons. Expectations	UMCSENTx	level

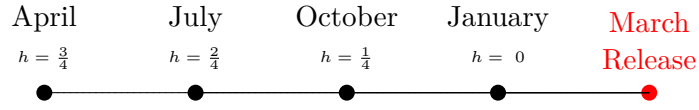
Notes: The symbols refer to the following transformations Δ: first difference, %Δ: growth rate.

In the nowcasting exercise our goal is to predict the yearly income distribution using timely information derived from quarterly indicators available during and after the year’s end and before the official release in March. For each year, we compute four nowcasts based on the different information sets given by the information available up to the release of the first, second, third and fourth quarterly data. The perspective is to consider at each update only the information available up to that point in time, so as to reproduce the nowcasts that would have been made in real-time. Our nowcasts origins are therefore in the months of April, July, October and January corresponding respectively to $h = \frac{3}{4}$, $h = \frac{2}{4}$, $h = \frac{1}{4}$ and $h = 0$. Figure 2 summarizes the four updates.

10. The inverse hyperbolic sine transformation is given by

$$x = g(z) = \ln \left(z + \sqrt{z^2 + 1} \right) = \sinh^{-1}(z)$$

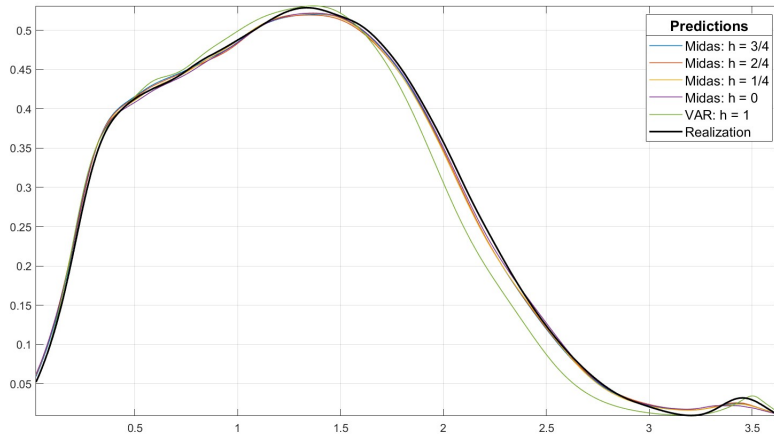
Figure 2: Updates in the nowcasting exercise



4.2 Results

The estimation sample begins in 1968, with predictions for income distribution starting from 1998 to 2023. For each period, we compute nowcasts at horizons $h = \frac{3}{4}$, $h = \frac{2}{4}$, $h = \frac{1}{4}$, and $h = 0$. Figure 3 presents a representative example of our forecasting output for the distribution of income in 1998. The figure compares the predictions generated by the functional MIDAS model in the four updates, alongside one-step-ahead forecasts from the functional VAR model, and the actual observed distribution.¹¹

Figure 3: Comparing nowcasts for 1998



Notes: The figure shows the nowcasts for the hyperbolic-sine transformed distribution of households' income normalized by GDP-per capita for the year 2002.

As our pseudo-real-time nowcasting exercise is designed to produce nowcasts of the income distribution for each year from 1998 to 2023 in four different updates, we are reporting the results in four distinct panels in table 3 and 4. Our competitors in the forecasting race are a simple functional VAR model for the low frequency distribution only and a functional MIDAS model with a ridge-type prior the coefficients in the SUR-MIDAS representation. We consider a ridge type prior both for an unrestricted MIDAS dynamics and for a restricted Almon-lag

¹¹. Also here, the point forecast of the distribution function is obtained by averaging across the posterior distribution draws.

dynamics.¹² The ridge-type prior assigns a distinct degree of shrinkage to the coefficients on the lags of the low frequency factors and the high frequency indicators. While the simple functional VAR model is updated only in April, when the release of the data on the survey of the previous year becomes available, the other models leverage high frequency information within the year to nowcast the low frequency distribution before its release. Overall, the Table shows that the functional MIDAS model does a pretty good job when compared to the simple functional single frequency VAR, with gains across all measures. In particular, including high frequency information allows to increase the accuracy of the forecast for the entire distribution function and some selected moments of this distribution. Notably, including high frequency information allows to reduced the RMSE associated to the two measures of inequality, being the Gini index and the coefficient of variation. This is true in all the updates considered. The spike-and-slab group lasso prior often outperforms the ridge prior in predicting the entire distribution function of income when comparing the KL distance with the true distribution; however, this superiority does not extend to all the metrics considered in this analysis.

12. Section [A.4](#) in the appendix provides the details on this prior and setting the hyper-parameters in this model.

Table 3: Forecast accuracy from 1998 to 2023

$h = \frac{3}{4}$							
	Avg KL	Avg HD	Avg QS5	Avg QS20	Avg QS50	Avg QS80	Avg QS95
FLAT VAR	0.1612	0.0845	0.0053	0.0114	0.0224	4407.8732	11962.0658
BVAR	0.0242	0.0441	0.0083	0.0151	0.0187	0.0119	126.7556
RIDGE-SUR-unres-MIDAS	0.0116	0.0349	0.0069	0.0154	0.0167	0.0110	0.0069
RIDGE-SUR-MIDAS	0.0114	0.0383	0.0086	0.0171	0.0181	0.0120	0.0086
BLSS-SUR MIDAS	0.0090	0.0371	0.0069	0.0096	0.0127	0.0093	0.0058
$h = \frac{2}{4}$							
	Avg KL	Avg HD	Avg QS5	Avg QS20	Avg QS50	Avg QS80	Avg QS95
FLAT VAR	0.1612	0.0845	0.0053	0.0114	0.0224	4407.8732	11962.0658
BVAR	0.0242	0.0441	0.0083	0.0151	0.0187	0.0119	126.7556
RIDGE-SUR-unres-MIDAS	0.0118	0.0349	0.0068	0.0149	0.0162	0.0108	0.0070
RIDGE-SUR-MIDAS	0.0115	0.0378	0.0086	0.0167	0.0177	0.0117	0.0083
BLSS-SUR MIDAS	0.0129	0.0376	0.0064	0.0106	0.0128	0.0088	0.0092
$h = \frac{1}{4}$							
	Avg KL	Avg HD	Avg QS5	Avg QS20	Avg QS50	Avg QS80	Avg QS95
FLAT VAR	0.1612	0.0845	0.0053	0.0114	0.0224	4407.8732	11962.0658
BVAR	0.0242	0.0441	0.0083	0.0151	0.0187	0.0119	126.7556
RIDGE-SUR-unres-MIDAS	0.0117	0.0350	0.0069	0.0152	0.0166	0.0109	0.0069
RIDGE-SUR-MIDAS	0.0105	0.0374	0.0086	0.0162	0.0172	0.0113	0.0075
BLSS-SUR MIDAS	0.0088	0.0374	0.0067	0.0118	0.0148	0.0097	0.0058
$h = 0$							
	Avg KL	Avg HD	Avg QS5	Avg QS20	Avg QS50	Avg QS80	Avg QS95
FLAT VAR	0.1612	0.0845	0.0053	0.0114	0.0224	4407.8732	11962.0658
BVAR	0.0242	0.0441	0.0083	0.0151	0.0187	0.0119	126.7556
RIDGE-SUR-unres-MIDAS	0.0117	0.0347	0.0068	0.0150	0.0163	0.0109	0.0069
RIDGE-SUR-MIDAS	0.0106	0.0373	0.0084	0.0161	0.0172	0.0114	0.0076
BLSS-SUR MIDAS	0.0109	0.0360	0.0066	0.0114	0.0146	0.0102	0.0079

Notes: Please see the text for the definitions of the various evaluation measures.

Table 4: Forecast accuracy metrics from 1998 to 2023 RMSE of selected moments

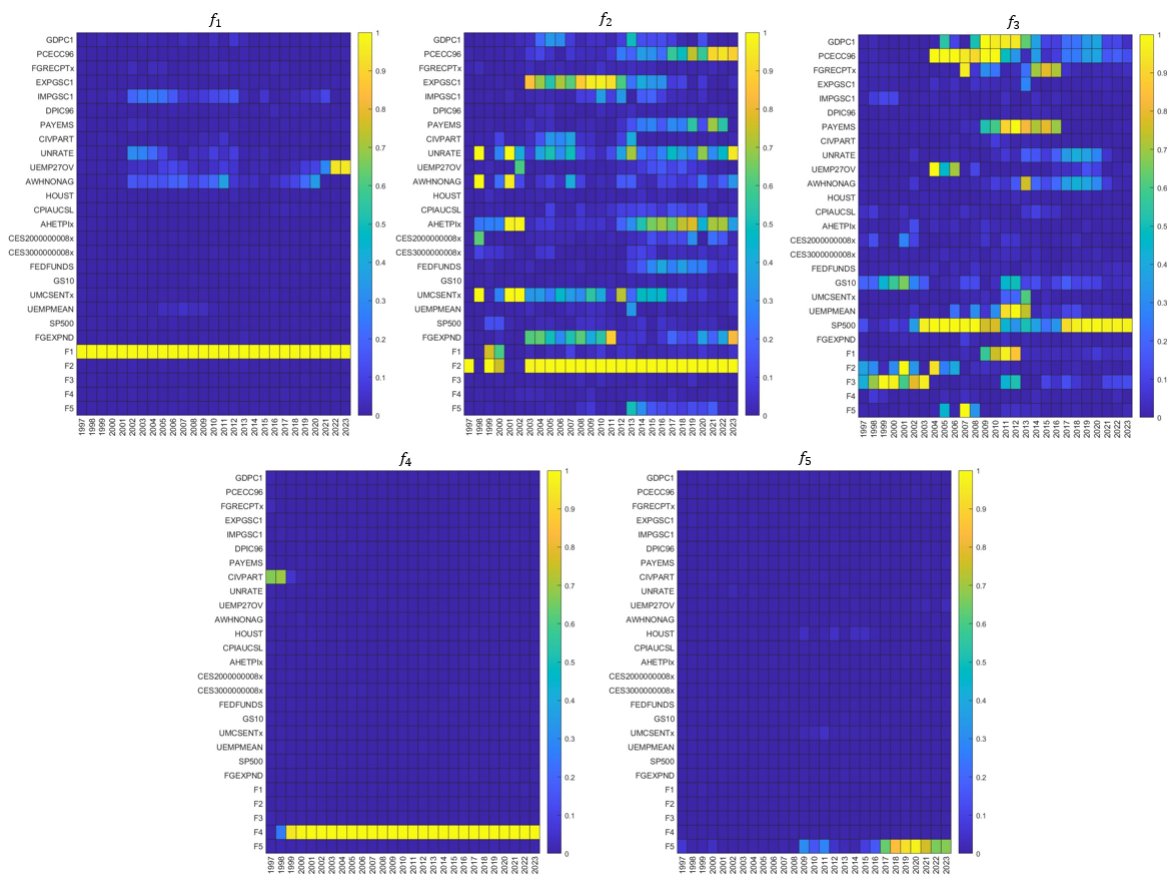
$h = \frac{3}{4}$							
	RMSE Mean	RMSE Variance	RMSE Skewness	RMSE Kurtosis	RMSE IR	RMSE GINI	RMSE CV
FLAT VAR	0.2932	0.0553	1.0749	3.3231	27196.4246	0.3993	16.1803
BVAR	0.0842	0.0432	0.2978	0.6516	0.0334	0.0677	2.3908
RIDGE-SUR-unres-MIDAS	0.0387	0.0357	0.1398	0.2913	0.0278	0.0207	1.3575
RIDGE-SUR-MIDAS	0.0480	0.0394	0.1537	0.3189	0.0274	0.0239	1.3073
BLSS-SUR MIDAS	0.0422	0.0388	0.1332	0.2882	0.0330	0.0228	1.3976
$h = \frac{2}{4}$							
	RMSE Mean	RMSE Variance	RMSE Skewness	RMSE Kurtosis	RMSE IR	RMSE GINI	RMSE CV
FLAT VAR	0.2932	0.0553	1.0749	3.3231	27196.4246	0.3993	16.1803
BVAR	0.0842	0.0432	0.2978	0.6516	0.0334	0.0677	2.3908
RIDGE-SUR-unres-MIDAS	0.0381	0.0356	0.1419	0.2953	0.0280	0.0212	1.3591
RIDGE-SUR-MIDAS	0.0460	0.0386	0.1535	0.3190	0.0277	0.0239	1.3099
BLSS-SUR MIDAS	0.0443	0.0372	0.1823	0.3958	0.0314	0.0303	1.6204
$h = \frac{1}{4}$							
	RMSE Mean	RMSE Variance	RMSE Skewness	RMSE Kurtosis	RMSE IR	RMSE GINI	RMSE CV
FLAT VAR	0.2932	0.0553	1.0749	3.3231	27196.4246	0.3993	16.1803
BVAR	0.0842	0.0432	0.2978	0.6516	0.0334	0.0677	2.3908
RIDGE-SUR-unres-MIDAS	0.0389	0.0358	0.1409	0.2932	0.0278	0.0209	1.3618
RIDGE-SUR-MIDAS	0.0445	0.0385	0.1456	0.3078	0.0278	0.0227	1.2497
BLSS-SUR MIDAS	0.0430	0.0390	0.1333	0.2916	0.0295	0.0192	1.4507
$h = 0$							
	RMSE Mean	RMSE Variance	RMSE Skewness	RMSE Kurtosis	RMSE IR	RMSE GINI	RMSE CV
FLAT VAR	0.2932	0.0553	1.0749	3.3231	27196.4246	0.3993	16.1803
BVAR	0.0842	0.0432	0.2978	0.6516	0.0334	0.0677	2.3908
RIDGE-SUR-unres-MIDAS	0.0378	0.0354	0.1406	0.2927	0.0280	0.0209	1.3631
RIDGE-SUR-MIDAS	0.0443	0.0388	0.1455	0.3057	0.0278	0.0223	1.2314
BLSS-SUR MIDAS	0.0420	0.0374	0.1632	0.3243	0.0311	0.0259	1.5719

Notes: Please see the text for the definitions of the various evaluation measures.

We can exploit the posterior estimates from the SUR-MIDAS model with spike and slab prior to investigate the set of relevant regressors, and understand whether the relevance of the predictors has changed over time. Figure 4 reports the posterior mean inclusion probabilities of the regressors in the five equations of the spike and slab lasso SUR-MIDAS model for the factors used for the approximation of the households' income distribution. The model is the one corresponding to $h = 0$, with high-frequency information fully matching the low frequency. The figure shows that some predictors have been historically relevant for predicting households' income distribution. More specifically, the lags of the unemployment rate, the number of civilians unemployed for 27 weeks and over, personal consumption expenditure, the number of non-farm employees, federal consumption expenditure and the SP500 are the most relevant predictors. Some groups, like those consisting of the lags of GDP and exports, were considered relevant and selected in certain periods, but their significance has diminished in recent years. Furthermore,

the groups corresponding to the own lags of the first, second and fourth factors are almost always selected and considered relevant, meaning that these factors exhibit a significant degree of persistence. The figure also shows that the first, fourth and fifth factors are predicted by a small, sparse set of predictors, indicating that only a limited number of predictors contribute significantly to their prediction. In contrast, the second and third factor exhibits less sparsity, meaning that a broader range of predictors are included and contribute to their prediction.

Figure 4: Inclusion probabilities in the SUR-MIDAS for $h = 0$



Notes: The heat-map reports the posterior mean probability of inclusion of each group.

5 Conclusions

In this paper, we propose a functional MIDAS model to leverage high-frequency information for nowcasting a distribution observed at a lower frequency. First, we approximate the low-frequency distribution using functional principal component analysis. Then, we specify a spike and slab prior for the parameters of the SUR-MIDAS approximation of the functional MIDAS

model. Through a simulation exercise, we demonstrate that this prior is particularly useful for selecting relevant predictors from a potentially large set of high-frequency indicators, even in small samples, as is typically the case when forecasting distributions from yearly household surveys. The simulation also highlights the importance of exploiting high-frequency indicators to accurately predict the entire distribution observed at low frequency, as well as key moments and features that can inform changes in inequality.

In general, although it does not enable high-frequency predictions for the micro-level distribution, the MIDAS framework is particularly promising for forecasting distributions by leveraging high-frequency information. It allows for imposing less structure when approximating the micro-level variable distribution, which would otherwise be required in the mixed-frequency framework of Schorfheide et al. 2015. For example, when the observed low-frequency variable of interest is a flow variable that accumulates high-frequency streams, aggregating the unobserved high-frequency distributions is non-trivial and may require parametric assumptions for tractability.

In the application, we aim nowcasting the March ASEC households' income data from the Current Population Survey leveraging on quarterly macroeconomic indicators available before the release. We find that our model enhances the accuracy of the forecasts for the distribution of households income and of key features of this distribution that signal changes in inequality. This enhanced predictive capability can help policymakers intervene more effectively with policies aimed at tackling inequality and promoting economic stability and growth.

References

- Aiyagari, S Rao. 1994. "Uninsured idiosyncratic risk and aggregate saving." *The Quarterly Journal of Economics* 109 (3): 659–684.
- Atchadé, Yves F., Gersende Fort, Eric Moulines, and Pierre Priouret. 2011. "Adaptive Markov chain Monte Carlo: Theory and methods." In *Bayesian Time Series Models*, edited by David Barber, Ali Taylan Cemgil, and Silvia Chiappa, 32–51. Cambridge, UK: Cambridge University Press.

- Azzalini, Adelchi, and Antonella Capitanio. 2003. “Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t-distribution.” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 65 (2): 367–389.
- Babii, Andrii, Luca Barbaglia, Eric Ghysels, and Jonas Striaukas. 2024. “Nowcasting and Aggregation: Why Small Euro Area Countries Matter.” *Available at SSRN 4902689*.
- Babii, Andrii, Eric Ghysels, and Jonas Striaukas. 2021. “Machine Learning Time Series Regressions With an Application to Nowcasting.” *Journal of Business & Economic Statistics* 0 (0): 1–23.
- Bayer, Christian, Ralph Luetticke, Lien Pham-Dao, and Volker Tjaden. 2019. “Precautionary savings, illiquid assets, and the aggregate consequences of shocks to household income risk.” *Econometrica* 87 (1): 255–290.
- Bilbiie, Florin O, Giorgio Primiceri, and Andrea Tambalotti. 2023. *Inequality and business cycles*. Technical report. National Bureau of Economic Research.
- Carriero, Andrea, Joshua Chan, Todd E. Clark, and Massimiliano Marcellino. 2022. “Corrigendum to “Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors” [J. Econometrics 212 (1) (2019) 137–154].” *Journal of Econometrics* 227 (2): 506–512.
- Carriero, Andrea, Todd Clark, and Massimiliano Marcellino. 2015. “Realtime nowcasting with a Bayesian mixed frequency model with stochastic volatility.” *Journal of the Royal Statistical Society Series A* 178 (4): 837–862.
- Carriero, Andrea, Todd E Clark, and Massimiliano Marcellino. 2019. “Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors.” *Journal of Econometrics* 212 (1): 137–154.
- Carter, Chris K., and Robert Kohn. 1994. “On Gibbs sampling for state space models.” *Biometrika* 81 (3): 541–553.
- Chan, Joshua C. C. 2022. “Asymmetric conjugate priors for large Bayesian VARs.” *Quantitative Economics* 13 (3): 1145–1169.

- Chan, Joshua CC, Liana Jacobi, and Dan Zhu. 2020. “Efficient selection of hyperparameters in large Bayesian VARs using automatic differentiation.” *Journal of Forecasting* 39 (6): 934–943.
- Chan, Joshua CC, and Ivan Jeliazkov. 2009. “Efficient simulation and integrated likelihood estimation in state space models.” *International Journal of Mathematical Modelling and Numerical Optimisation* 1 (1-2): 101–120.
- Chang, Minsu, Xiaohong Chen, and Frank Schorfheide. 2021. *Heterogeneity and aggregate fluctuations*. Technical report. National Bureau of Economic Research.
- Chang, Minsu, and Frank Schorfheide. 2022. “On the effects of monetary policy shocks on earnings and consumption heterogeneity.”
- De Polis, Andrea, Gary Koop, Stuart McIntyre, James Mitchell, and Ping Wu. 2024. “Mixed Frequency Functional VARs for Nowcasting and Structural Analysis of the Income Distribution in the UK.” Manuscript in preparation.
- Ghysels, Eric, Pedro Santa-Clara, and Rossen Valkanov. 2004. *The MIDAS Touch: Mixed Data Sampling Regression Models*. Technical report. University of California at Los Angeles, Anderson Graduate School of Management.
- Ghysels, Eric, Ross Valkanov, and Massimiliano Marcellino. 2025. *The Econometric Analysis of Mixed Frequency Data*. Forthcoming. Cambridge University Press.
- Huber, Florian, Massimiliano Marcellino, and Tommaso Tornese. 2024. “The Distributional Effects of Economic Uncertainty.” Manuscript.
- Jones, M.C. 1992. “Estimating densities, quantiles, quantile densities and density quantiles.” *Annals of the Institute of Statistical Mathematics* 44 (4): 721–727.
- Kaplan, Greg, Benjamin Moll, and Giovanni L Violante. 2018. “Monetary policy according to HANK.” *American Economic Review* 108 (3): 697–743.
- Kokoszka, Piotr, Hong Miao, Alexander Petersen, and Han Lin Shang. 2019. “Forecasting of density functions with an application to cross-sectional and intraday returns.” *International Journal of Forecasting* 35 (4): 1304–1317.

- Mogliani, Matteo, and Anna Simoni. 2021. “Bayesian MIDAS penalized regressions: Estimation, selection, and prediction.” *Journal of Econometrics* 222 (1, Part C): 833–860.
- Parzen, Emanuel. 1979. “Nonparametric Statistical Data Modeling.” *Journal of the American Statistical Association* 74 (365): 105–131.
- Petersen, Alexander, and Hans-Georg Müller. 2016. “Functional data analysis for density functions by transformation to a Hilbert space.” *The Annals of Statistics* 44 (1): 183–218.
- Ramsay, J. O., and B. W. Silverman. 2005. *Functional Data Analysis*. 2nd ed. Springer Series in Statistics. Springer-Verlag New York.
- Schorfheide, Frank, and Dongho Song. 2015. “Real-time forecasting with a mixed-frequency VAR.” *Journal of Business & Economic Statistics* 33 (3): 366–380.
- Xu, Xiaofan, and Malay Ghosh. 2015. “Bayesian Variable Selection and Estimation for Group Lasso.” *Bayesian Analysis* 10 (4): 909–936.
- Zellner, Arnold, and Tomohiro Ando. 2010. “A direct Monte Carlo approach for Bayesian analysis of the seemingly unrelated regression model.” *Journal of Econometrics* 159, no. 1 (November): 33–45.

A Appendix

A.1 SUR-MIDAS approximation of the functional MIDAS model

The functional MIDAS model is given by:

$$q_t(\tau) = c_q(\tau) + \sum_{l=0}^{p_x-1} \mathbf{B}_{qx,l}(\tau) L(l/m) \mathbf{x}_t^{(m)} + \sum_{l=1}^{p_q} \int_0^1 B_{qq,l}(\tau, \tau') L(l) q_t(\tau') d\tau' + u_{q,t}(\tau) \quad (26)$$

Thanks to the finite dimensional approximation (4) and assuming $\mu(\tau) = 0$ for ease of exposition, we can rewrite the function MIDAS expanding the unknown functions of τ as follows

$$\begin{aligned} c_q(\tau) &= \mathbf{h}_K(\tau)' \tilde{c}_{q,t} & \mathbf{B}_{qx,l}(\tau) &= \mathbf{h}_K(\tau)' \mathbf{B}_{qx,l} \\ B_{qq,l}(\tau, \tau') &= \mathbf{h}_K(\tau)' \mathbf{B}_{qq,l} \mathbf{s}_K(\tau') & \mathbf{u}(\tau) &= \mathbf{h}_K(\tau)' \tilde{u}_{q,t} \end{aligned} \quad (27)$$

where $\mathbf{s}_K(\tau')$ is a K -dimensional vector of functional basis such that

$$\int \mathbf{s}_K(\tau) \mathbf{h}_K(\tau)' d\tau = \mathbf{C}_f \quad (28)$$

Plugging in (26) we get

$$\mathbf{h}_K(\tau)' \mathbf{f}_{t;K} = \mathbf{h}_K(\tau)' \tilde{c}_{q,t} + \sum_{l=1}^{p_x} \mathbf{h}_K(\tau)' \mathbf{B}_{qx,l} L(l/m) \mathbf{x}_t^{(m)} + \mathbf{h}_K(\tau)' \sum_{l=1}^{p_q} \mathbf{B}_{qq,l} L(l) \mathbf{f}_{t;K} + \mathbf{h}_K(\tau)' \tilde{u}_{q,t} \quad (29)$$

that becomes

$$\mathbf{f}_{t;K} = \tilde{c}_{q,t} + \sum_{l=1}^{p_x} \mathbf{B}_{qx,l} L(l/m) \mathbf{x}_t^{(m)} + \sum_{l=1}^{p_q} \mathbf{B}_{l,qq} \mathbf{C}_f L(l) \mathbf{f}_{t;K} + \tilde{u}_{q,t} \quad (30)$$

Now define $\Phi_{fx,l}$, $\Phi_{ff,l}$ and $\mathbf{u}_{f,t+h} = \tilde{u}_{q,t}$ and we obtain the SUR MIDAS model

$$\mathbf{f}_{t;K} = \phi_0 + \sum_{l=1}^{p_x} \Phi_{fx,l} L(l/m) \mathbf{x}_t^{(m)} + \sum_{l=1}^{p_q} \Phi_{ff,l} L(l) \mathbf{f}_{t;K} + \mathbf{u}_{f,t} \quad (31)$$

A.2 Conditional posterior distributions in the Gibbs Sampler

Exploiting the direct representation of the Almon lag polynomial we can rewrite the SUR-MIDAS model as

$$\begin{aligned}
f_{1t} &= \mathbf{z}'_t \boldsymbol{\theta}_1 + \varepsilon_{1t} & \varepsilon_{1t} &\sim \mathcal{N}(0, \sigma_1^2) \\
&\vdots & & \\
f_{2t} &= \mathbf{z}'_t \boldsymbol{\theta}_2 + \alpha_{21} u_{1t} + \varepsilon_{2t} & \varepsilon_{2t} &\sim \mathcal{N}(0, \sigma_2^2) \\
&\vdots & & \\
f_{Kt} &= \mathbf{z}'_t \boldsymbol{\theta}_K + \alpha_{K,1} u_{1t} + \dots + \alpha_{K,K-1} u_{K-1t} + \varepsilon_{Kt} & \varepsilon_{Kt} &\sim \mathcal{N}(0, \sigma_K^2)
\end{aligned} \tag{32}$$

where the vector $\boldsymbol{\theta}_i = [\boldsymbol{\theta}_{i,1}', \dots, \boldsymbol{\theta}_{i,G}']'$ collects all the parameters of the Almon lag polynomial for all the $G = (n_x + K)$ groups in the i^{th} equation for $i = 1, \dots, K$. Note that each vector $\boldsymbol{\theta}_{ij}$ is of dimension $(p_a + 1 - r_a) \times 1$. Define \mathbf{F} as the $T \times K$ matrix with the $T \times 1$ vector $\mathbf{f}_i = [f_{i1}, \dots, f_{iT}]'$ stacked on its columns and define $\boldsymbol{\Theta}$ the $G(p_a + 1 - r_a) \times K$ matrix stacking the $\boldsymbol{\theta}_i$ on its columns. In compact form the model is

$$(\mathbf{F} - \mathbf{Z}\boldsymbol{\Theta})\tilde{\mathbf{A}}' = \boldsymbol{\varepsilon} \tag{33}$$

where $\tilde{\mathbf{A}}$ is the lower triangular matrix with unit elements on the main diagonal and rows given by $\tilde{\boldsymbol{\alpha}}_{i,\cdot} = [-\alpha_{i,1}, \dots, -\alpha_{ii-1}]$. Exploiting the triangularization in Carriero et al. (2022) to rewrite the equations that involve the coefficients in the i^{th} equation i.e. $\boldsymbol{\theta}_i$ as

$$vec((\mathbf{F} - \mathbf{Z}\boldsymbol{\Theta}^{[i=0]})\tilde{\mathbf{A}}^{(i:K,1:K)'}) = (\tilde{\mathbf{A}}^{(i:K,i)} \otimes \mathbf{Z}) \boldsymbol{\theta}_i + vec(\boldsymbol{\varepsilon}^{(i:K)}) \tag{34}$$

where the notation $\boldsymbol{\Theta}^{[i=0]}$ denotes the matrix $\boldsymbol{\Theta}$ with the i^{th} column set equal to zero, while the notation $\tilde{\mathbf{A}}^{(i:K,i)}$ denotes the elements from row i to row K in the i^{th} column of $\tilde{\mathbf{A}}$. Define

$$\tilde{\mathbf{Z}} = (\tilde{\mathbf{A}}^{(j:K,j)} \otimes \mathbf{Z}) \tag{35}$$

and $\tilde{\mathbf{Z}}_j$ the j^{th} column of $\tilde{\mathbf{Z}}$ while $\tilde{\mathbf{Z}}_{\setminus j}$ the matrix $\tilde{\mathbf{Z}}$ without the j^{th} column. Now rewrite it as

$$vec((\mathbf{F} - \mathbf{Z}\boldsymbol{\Theta}^{[i=0]})\tilde{\mathbf{A}}^{(i:K,1:K)'}) = \tilde{\mathbf{Z}}_j \boldsymbol{\theta}_{ij} + \tilde{\mathbf{Z}}_{\setminus j} \boldsymbol{\theta}_{i\setminus j} + vec(\boldsymbol{\varepsilon}^{(i:K)}) \tag{36}$$

then consider

$$vec((\mathbf{F} - \mathbf{Z}\Theta^{[i=0]})\tilde{\mathbf{A}}^{(i:K,1:K)'}) - \tilde{\mathbf{Z}}_{\setminus j}\boldsymbol{\theta}_{i\setminus j} = \tilde{\mathbf{Z}}_j\boldsymbol{\theta}_{ij} + vec(\boldsymbol{\varepsilon}^{(i:K)}) \quad (37)$$

Defining $\tilde{\mathbf{f}}_{ij} = vec((\mathbf{F} - \mathbf{Z}\Theta^{[i=0]})\tilde{\mathbf{A}}^{(i:K,1:K)'}) - \tilde{\mathbf{Z}}_{\setminus j}\boldsymbol{\theta}_{i\setminus j}$ becomes

$$\tilde{\mathbf{f}}_{ij} = \tilde{\mathbf{Z}}_j\boldsymbol{\theta}_{ij} + vec(\boldsymbol{\varepsilon}^{(i:K)}) \quad (38)$$

To obtain the conditional posterior of the $\boldsymbol{\theta}_{ij}$'s just combine this equation with the prior for $\boldsymbol{\theta}_{ij}$ in equation (18). In particular we get

$$\begin{aligned} \boldsymbol{\theta}_{ij} \mid \boldsymbol{\theta}_{i\setminus j}, \boldsymbol{\theta}_{1:i-1}, \tau_{ij}, \lambda, \pi_{1i}, \pi_{0i}, \mathbf{f}, \mathbf{Z} &\sim (1 - \pi_{1ij})\mathcal{N}(\mathbf{P}_{ij}^{-1}\mathbf{C}_{ij}, \sigma_i^2\mathbf{P}_{ij}^{-1}) + \pi_{1ij}\delta_0(\boldsymbol{\theta}_{ij}) \\ \mathbf{P}_{ij} &= \tau_{ij}^2\mathbf{I}_{gg} + \tilde{\mathbf{Z}}_j'\boldsymbol{\Sigma}^{(i:K)^{-1}}\tilde{\mathbf{Z}}_j \\ \mathbf{C}_{ij} &= \tilde{\mathbf{Z}}_j'\boldsymbol{\Sigma}^{(i:K)^{-1}}\tilde{\mathbf{f}}_{ij} \end{aligned} \quad (39)$$

To write the conditional posterior of the elements in $\tilde{\mathbf{A}}$ we define

$$\mathbf{u}_i = \mathbf{f}_i - \mathbf{Z}\boldsymbol{\theta}_i \quad (40)$$

where \mathbf{f}_i is $T \times 1$ and \mathbf{Z} is the $T \times (p_a + 1 - r_a)G$ design matrix. Furthermore for $i = 2, \dots, K$ define $\mathbf{S}_i = diag(s_1^2, \dots, s_i^2)$. Define also

$$\tilde{\mathbf{U}}_i = [\mathbf{u}_1, \dots, \mathbf{u}_{i-1}]' \quad \text{for } 2 = 1, \dots, K \quad (41)$$

and \mathbf{A} the lower triangular matrix with unit elements on the main diagonal, such that the free elements in each row i are given by $\boldsymbol{\alpha}_{i,\cdot} = [\alpha_{i,1}, \dots, \alpha_{ii-1}]$ for $i = 1, \dots, K$.

$$\boldsymbol{\alpha}_{i,\cdot}' \sim \mathcal{N}\left(\left(\tilde{\mathbf{U}}_i'\tilde{\mathbf{U}}_i + \mathbf{S}_i^{-1}\right)^{-1}\left(\tilde{\mathbf{U}}_i'\mathbf{u}_i\right), \left(\tilde{\mathbf{U}}_i'\tilde{\mathbf{U}}_i + \mathbf{S}_i^{-1}\right)^{-1}\sigma_i^2\right) \quad \text{for } 2 = 1, \dots, K \quad (42)$$

The other conditional posteriors are given by

$$\begin{aligned} \sigma_i^2 \mid \boldsymbol{\theta}, \tau_{ij}, \lambda_{ij}, \pi_{1i}, \pi_{0i}, \mathbf{f}, \mathbf{Z} &\sim \text{Inv-Gamma}\left(\frac{T + \tilde{G}_i - 1}{2} + \frac{(i-1)}{2} + \frac{(v_{0i} + i - K)}{2}, \right. \\ &\left. \frac{1}{2}\|(\mathbf{Y} - \mathbf{Z}\Theta)\tilde{\mathbf{A}}'\|_2^2 + \frac{1}{2}\sum_{j=1}^G \frac{\|\boldsymbol{\theta}_{ij}\|_2^2}{\tau_{ij}^2} + \frac{s_i^2}{2}\right) \end{aligned} \quad (43)$$

$$\tau_{ij}^{-2} \mid \boldsymbol{\theta}, \sigma_i^2, \lambda_{ij}, \pi_{1i}, \pi_{0i}, \mathbf{f}, \mathbf{Z} \sim \begin{cases} \text{Inv-Gaussian} \left(\frac{\lambda_{ij}\sigma_i}{\|\boldsymbol{\theta}_{ij}\|_2}, \lambda_{ij}^2 \right) & \text{if } \boldsymbol{\theta}_{ij} \neq 0 \\ \text{Gamma} \left(\frac{g_{ij}+1}{2}, \frac{\lambda_{ij}^2}{2} \right) & \text{if } \boldsymbol{\theta}_{ij} = 0 \end{cases} \quad (44)$$

$$\pi_{0i} \mid \boldsymbol{\theta}, \sigma_i^2, \tau_{ij}, \lambda, \pi_{1i}, \mathbf{f}, \mathbf{Z} \sim \text{Beta} \left(\sum_{j=1}^G (1 - \gamma_{ij}) + c, \sum_{j=1}^G \gamma_{ij} + d \right) \quad (45)$$

where v_{i0} is set equal to $1 + \frac{i}{2}$, $\tilde{G}_i = \sum_{j=1}^G g_{ij} \gamma_{ij}$, $d = 1$, $c = \bar{\kappa} G^\nu$, $\bar{\kappa} = (1 + G^{-1})$ and

$$\gamma_{ij} = \begin{cases} 1 & \text{if } \boldsymbol{\theta}_{ij} \neq 0 \\ 0 & \text{if } \boldsymbol{\theta}_{ij} = 0 \end{cases} \quad (46)$$

$$\pi_{1,ij} = P(\boldsymbol{\theta}_{ij} = 0 \mid \theta_{\setminus j}, \sigma^2, \tau, \lambda, \pi_{0i}, \mathbf{f}, \mathbf{Z}) = \frac{\pi_{0i}}{\pi_{0i} + (1 - \pi_{0i}) \left[(\tau_j^2)^{-\frac{g_j}{2}} |\mathbf{P}_{ij}|^{-\frac{1}{2}} \exp \left(\frac{1}{2\sigma^2} \mathbf{C}'_{ij} \mathbf{P}_{ij}^{-1} \mathbf{C}_{ij} \right) \right]} \quad (47)$$

Assuming a conjugate Gamma prior on the penalty hyper-parameters λ_{ij}^2 parameterized by shape parameter a_2 and rate parameter b_2 we get the following conditional posterior

$$\lambda_{ij}^2 \mid \boldsymbol{\theta}, \sigma_i^2, \tau_{ij}, \pi_{1i}, \pi_{0i}, \mathbf{f}, \mathbf{Z} \sim \text{Gamma} \left(\frac{g_{ij} + 1}{2} + a_2, \frac{\tau_{ij}^2}{2} + b_2 \right) \quad (48)$$

A.3 Monte Carlo simulation

Here we report the details on the parameters of the Monte Carlo exercise in section 3. We mostly built on the Monte Carlo exercise in Mogliani et al. (2021) for what concerns setting the value of the parameters of the DGP, extending it to a SUR framework. More in the specific, we set $\boldsymbol{\Sigma}_\epsilon$ such that the correlation between $x_{j,t}^{(m)}$ and $x_{j',t}^{(m)}$ with $j' \neq j$ is 0.50. In equation (24) we

set $\mu = 0.1$ and $\rho = 0.5$. In equation (23) and we set $\alpha_i = 0$ and

$$\beta_1 = \begin{pmatrix} 0 \\ 0.3 \\ 0.3 \\ 0 \\ -0.5 \\ 0.1 \\ 0 \\ 0 \\ 0.3 \\ \mathbf{0}_{21 \times 1} \end{pmatrix} \quad \beta_2 = \begin{pmatrix} \mathbf{0}_{19 \times 1} \\ 0.3 \\ 0.3 \\ 0.1 \\ 0 \\ -0.5 \\ -0.3 \\ 0 \\ 0.3 \\ 0.1 \\ \mathbf{0}_{2 \times 1} \end{pmatrix} \quad \beta_3 = \begin{pmatrix} 0 \\ 0.3 \\ 0.5 \\ 0 \\ -0.1 \\ \mathbf{0}_{21 \times 1} \\ 0.3 \\ 0 \\ 0 \\ 0.5 \end{pmatrix} \quad \Phi_1 = \begin{bmatrix} 0.8 & 0.001 & 0 \\ 0.02 & 0.5 & 0 \\ -0.01 & 0 & 0.5 \end{bmatrix} \quad \Phi_2 = \begin{bmatrix} 0.02 & 0 & 0.1 \\ 0 & 0.03 & 0.04 \\ 0 & -0.02 & 0.05 \end{bmatrix} \quad (49)$$

Conditional on these parameters, we set the diagonal elements of Ω such that the noise-to-signal ratio of the mixed-frequency regression is 0.20. The off diagonal elements of Ω are instead set to imply $\text{corr}(u_{1,t}, u_{2,t}) = 0.1$, $\text{corr}(u_{1,t}, u_{3,t}) = -0.1$ and $\text{corr}(u_{2,t}, u_{3,t}) = 0.2$.

A.4 Ridge type shrinkage prior for the unrestricted SUR-MIDAS

We compare the lasso-type shrinkage prior with a ridge-type shrinkage prior for the coefficients in the SUR-MIDAS model. This prior is inspired by the shrinkage prior for a univariate unrestricted MIDAS regression in Carriero et al. (2015). More specifically, we rewrite the SUR-MIDAS model in compact form as

$$\mathbf{F} = \mathbf{X}\Phi + \mathbf{U} \quad (50)$$

where \mathbf{F} is the $T \times K$ matrix of factors, \mathbf{X} is the $T \times (n_x p_x + K p_q)$ design matrix of regressors, which contains both the lags of the high frequency macroeconomic indicators and the lags of the low frequency factors. \mathbf{U} is a $T \times N$ matrix of correlated errors such that $E[\mathbf{U}'\mathbf{U}] = \Omega$. We specify the following prior for (Φ, Ω)

$$\text{vec}(\Phi) | \Omega \sim \mathcal{N}(\underline{\Phi}, \underline{V}) \quad (51)$$

$$\Omega \sim \mathcal{IW}(\mathcal{S}, v_0) \quad (52)$$

where $\underline{\Phi}$ is centering the coefficient of the first own lag of the low frequency factors on 0.8 and all the other coefficients to 0. Then the variance covariance matrix of the prior distribution is

given by

$$\underline{\mathbf{V}} = \underline{\mathbf{\Omega}} \otimes \underline{\mathbf{V}}_0 \quad (53)$$

where the elements of the diagonal matrix $\underline{\mathbf{V}}_0$ are

$$\omega_i = \begin{cases} \theta_1 & \text{intercept} \\ \frac{\theta_2}{s_r^{2l\theta^3}} & \text{on the lags of the low frequency factors} \\ \frac{\theta_4}{s_r^{2l\theta^5}} & \text{on the lags of the high frequency variables} \end{cases} \quad (54)$$

The shrinkage hyper-parameters $[\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]$ are selected to maximize the marginal data density of the model which is available in close form. We resort to automatic differentiation to perform the maximization of the marginal likelihood with respect to these hyper-parameters following Chan et al. (2020). For the model with the restricted Almon-lag dynamics we consider the same prior with $l = 1$.